McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

7 F - GEN

www.MHEST.com

F

Fabales — Fuzzy-structure acoustics

Fabales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of three closely related families (Mimosaceae, Caesalpiniaceae, Fabaceae) collectively called the legumes. Members of the order typically have stipulate, compound leaves, 10–many stamens which are often united by the filaments, and a single carpel which gives rise to a dry fruit (legume) that opens at maturity by splitting along two sutures, releasing the non-endospermous seeds. Many, or perhaps most, members of the order harbor symbiotic nitrogen-fixing bacteria in the roots.

Mimosaceae includes some 40 genera and 2000 species, most of which are woody and occur in the warmer regions of the world. The flowers are regular, and the stamens at least twice as many as the petals, and often very numerous. Common members of the family are *Acacia, Albizia, Mimosa,* and *Prosopis* (mesquite).

Caesalpiniaceae includes about 150 genera and 2200 species, most of which are woody and of tropical or subtropical regions. The flowers are somewhat irregular, typically with five (or fewer) unequal petals and 10 or fewer stamens. The family includes the redbud (*Cercis*), Kentucky coffee tree (*Gymnocladus*), honey locust (*Gleditsia*), and numerous tropical ornamentals.

Fabaceae (sometimes termed the Papilionoideae) has about 400 genera and at least 10,000 species of trees, shrubs, vines, and herbs, which are widespread in the tropical and temperate regions. The flowers are characteristically irregular, with a large upper petal (the banner or standard), two wing petals, and the lower two petals fused to form a boat-shaped keel. Usually there are 10 stamens with 9 of them united by their filaments, and the tenth one more or less separate. Well-known members of the family include beans (*Phaseolus*), peas (*Pisum*), sweet peas (*Lathyrus*), clover (*Trifolium*), vetch (*Vicia*), alfalfa (*Medicago*), and lupine (*Lupinus*). *See* ALFALFA; BEAN; CLOVER; COWPEA; KENTUCKY COFFEE TREE; KUDZU; LESPEDEZA; LOCUST (FORESTRY); LUPINE; PEA; PEANUT; SOYBEAN; VETCH. T. M. Barkley

Facies (geology)

Any observable attribute of rocks, such as overall appearance, composition, or conditions of formation, and changes that may occur in these attributes over a geographic area. The term "facies" is widely used in connection with sedimentary rock bodies, but is not restricted to them. In general, facies are not defined for sedimentary rocks by features produced during weathering, metamorphism, or structural disturbance. In metamorphic rocks specifically, however, facies may be identified by the presence of minerals that denote degrees of metamorphic change.

Sedimentary Facies

The term "sedimentary facies" is applied to bodies of sedimentary rock on the basis of descriptive or interpretive characteristics. Descriptive facies are based on lithologic features such as composition, grain size, bedding characteristics, and sedimentary structures (lithofacies); on biological (fossil) components (biofacies); or on both. Individual lithofacies or biofacies may be single beds a few millimeters thick or a succession of beds tens to hundreds of meters thick. For example, a river deposit may consist of decimeter-thick beds of a conglomerate lithofacies interbedded with a cross-bedded sandstone lithofacies. The fill of certain major Paleozoic basins may be divided into units hundreds of meters thick comprising a shelly facies, containing such fossils as brachiopods and trilobites, and graptolitic facies. The scale of an individual lithofacies or biofacies unit depends on the level of detail incorporated in its definition, and is determined only by criteria of convenience, or

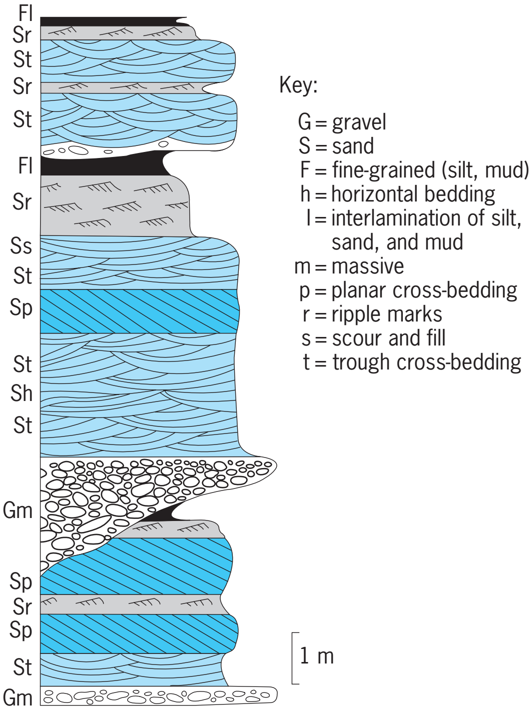


Fig. 1. Stratigraphic section through a fluvial deposit illustrating its subdivision into lithofacies and the use of a lithofacies code scheme. 1 m = 3.3 ft. (*After A. D. Miall, A review of the braided river depositional environment*, *Earth Sci. Rev., 13:1–62, 1977*)

availability of outcrop and subsurface data, depending on whether a geologist wishes to describe a single outcrop or an entire sedimentary basin.

Facies analysis. The term "facies" can be used in an interpretive sense for groups of rocks that are thought to have been formed under similar conditions. This usage may emphasize specific depositional processes such as a turbidite facies, or a particular depositional environment such as a shelf carbonate facies, encompassing a range of depositional processes. *See* TURBIDITE.

Facies analysis is an important component of stratigraphic and sedimentary geology because it provides a methodology of systematic description that is fundamental to orderly stratigraphic documentation and to interpretations of depositional environment and paleogeography. Most sedimentological basin studies should now begin with such an analysis, in which the succession is divided into its component lithofacies and their characteristics described carefully. Lithofacies may be identified by name or by a code system devised by the researcher, the latter method being convenient for abbreviated description and notetaking and for computer processing (**Fig. 1**).

Increasing use is being made of three-dimensional geometric (architectural) information regarding the extent of facies units, because such information commonly is environmentally diagnostic.

Mappable stratigraphic units commonly consist of a single lithofacies or a small assemblage of lithofacies. This provides the basis for a hierarchical subdivision of stratigraphic sequences ranging from the individual bed or lithofacies unit through member or lithosome and the higher-ranking subdivisions such as stratigraphic sequence, formation, and group. *See* FACIES (GEOLOGY); STRATIGRAPHY.

Facies assemblages. Groups of facies (usually lithofacies) that are commonly found together in the sedimentary record are known as facies assemblages or facies associations. These groupings provide the basis for defining broader, interpretive facies for the purpose of paleogeographic reconstruction. For example, a delta plain facies may be defined, consisting of distributary channel sandstones, levee sandstones and siltstones, and backswamp mudstones and coal. This may pass seaward into a delta front facies consisting of distributary mouth bar sandstones and interdistributary bay mudstones and shell beds. Still further seaward this may pass into a fine-grained prodelta facies, and landward there may be a fluvial facies and a lacustrine facies. An example of a carbonate facies assemblage is illustrated in **Fig. 2**.

The shelly facies and graptolitic facies described earlier are biofacies terms, but they correspond to distinct lithofacies assemblages dominated by shelf carbonate sediments and by deep-water mudstones and siltstones, respectively. Use of the terms provides a convenient way to describe the broad paleogeography of some of the major Paleozoic sedimentary basins around the margins of North America and elsewhere. *See* PALEOGEOGRAPHY.

Many facies assemblages are cyclic in the sense that the individual lithofacies components tend to follow each other in a similar order in a stratigraphic succession, reflecting the repeated occurrence of the same depositional processes. For example, coal beds commonly (but not invariably) follow a seat

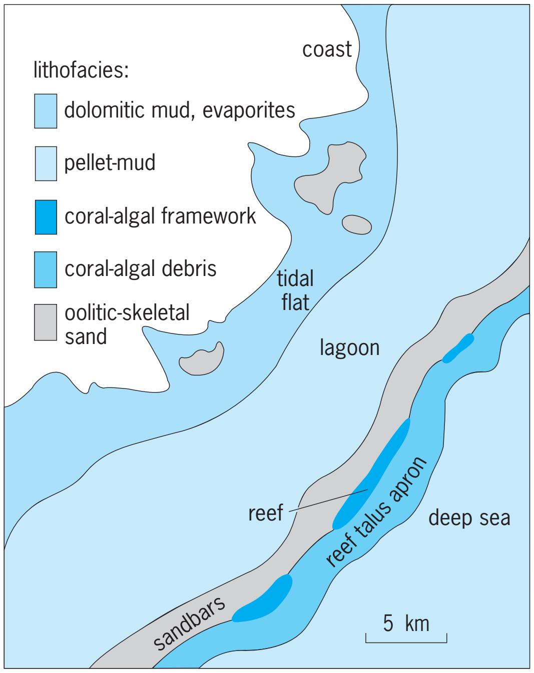


Fig. 2. Hypothetical carbonate shelf showing relationship of lithofacies to environmental facies; based on the modern shelves of Florida and the Bahamas. 1 km = 0.6 mi.

earth or root bed. Such assemblages are termed cyclothems or, simply, cycles. Much sedimentological research has been devoted to analysis and interpretation of cycles, which occur in most depositional settings. For example, the Carboniferous coal-bearing deposits of Europe and of central and eastern North America consitute the original cyclothems. Problems have arisen because these and other cycles can originate from several causes, the effects of which may be very similar and may be superimposed. Thus, cyclothems are caused partly by the advance and abandonment of delta lobes and partly by oscillatory changes of sea level. Second, progressive changes in a sedimentary basin, such as the slow seaward progradation of a delta or tidal flat, may be interrupted by rare random events, for example, a hurricane, which can erode much of the previous cyclic sedimentary record and leave its own unique lithofacies instead. Analysis of cyclicity is based on application of Walther's law. This law, first defined in the nineteenth century, stemmed from the recognition that vertical successions of strata are formed largely by lateral migration of sedimentary environments. The law states that only those facies found side by side in nature can occur superimposed on each other in vertical succession. Recognition of widespread cyclic units is the basis of sequence stratigraphy. Sequences record regional or global changes in sea level brought about by tectonics or eustacy. *See* CYCLOTHEM; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

Facies models. These have been defined for most sedimentary environments. Each model is intended as a summary of the depositional processes and products occurring in a specific setting, such as a tidal flat or a submarine fan. Such models are useful guides when carrying out preliminary interpretations of an unknown sedimentary succession, because they provide a range of criteria relating to lithofacies, sedimentary structures, biofacies, and cyclic characteristics that the geologist can use to confirm or negate a particular interpretation.

Facies maps. These are constructed to illustrate lithologic variations across a sedimentary basin, and they usually reveal patterns or trends that can readily be interpreted in terms of depositional environments and paleogeography. Various quantitative techniques have been devised for emphasizing relevant features. Ratio maps are particularly useful, for example, ones showing the ratio of total thickness of sandstone beds to total shale or total carbonate thickness. In elastic coastal environments such a map may reveal the location and shape of major sandstone bodies, which can then, from their size, shape, and orientation, be interpreted in terms of an appropriate sedimentary model (delta, barrier island system, submarine fan, and so on) in conjunction with diagnostic sedimentological criteria. Location and trend may be of considerable economic importance if, for example, porous sandstone bodies are known to contain pools of oil or gas. *See* BARRIER ISLANDS; PETROLEUM GEOLOGY.

Another technique is to plot the percentage of sandstone or aggregate sandstone thickness in a section. **Figure 3** illustrates a map of this type, and reveals a hitherto unsuspected pattern of lobate sandstone bodies containing a readily identifiable deltaic shape. Thicker sandstone beds lie closer to the sediment source, and so in this example transport was clearly from northwest to southeast.

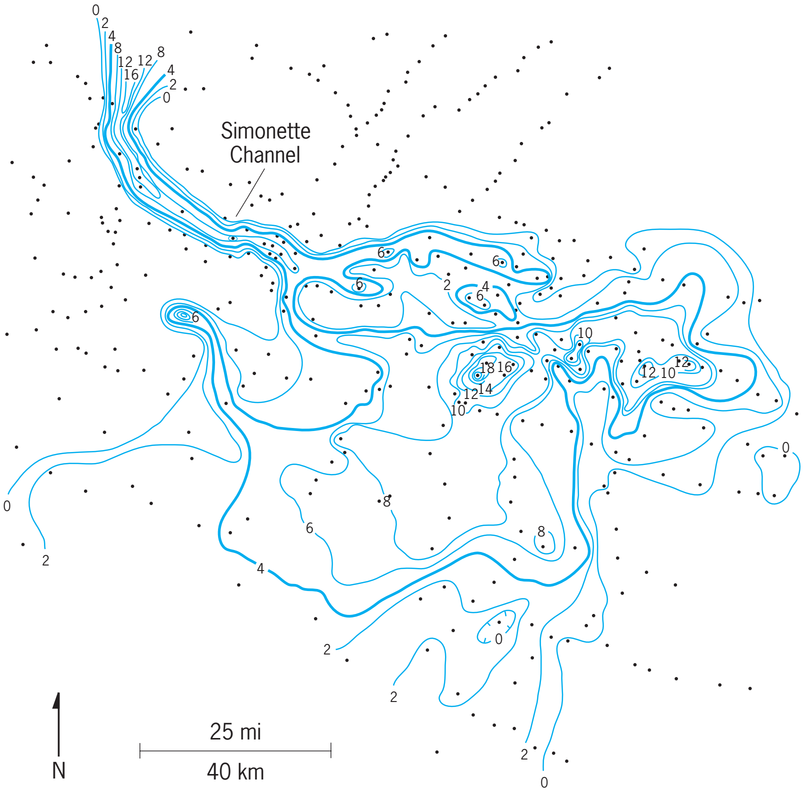


Fig. 3. Plot of sandstone thickness, in meters, in a Cretaceous sandstone-shale unit, derived from subsurface log analysis. Note the narrow sandstone-filled channel representing a delta distributary channel entering the area from the northwest, and the broad lobe of sandstone formed at the mouth of this channel where the delta splayed out into the sea. (*After J. Bhattacharya, Regional to subregional facies architecture of river-dominated deltas in the Alberta subsurface, Upper Cretaceous Dunvegan Formation, in Society for Sedimentary Geology, Concepts in Sedimentology and Paleontology, vol. 3, 1991*)

Similar techniques can be applied to many successions of carbonate rocks. Thus carbonates are typically classified in terms of the ratio of biogenic or detrital fragments to carbonate and matrix. This classification reflects depositional energy (wave and current strength) because the mud matrix tends to be winnowed out in high-energy environments. Therefore, a plot of the ratio can be a useful paleogeographic indicator. However, carbonates are particularly susceptible to diagenetic alteration, which can obscure primary depositional characteristics and hinder detailed paleogeographic reconstruction. *See* DIAGENESIS; SEDIMENTARY ROCKS; SEDIMENTOLOGY; STREAM TRANSPORT AND DEPOSITION.

Andrew D. Miall

Metamorphic Facies

A metamorphic facies is a collection of rocks containing characteristic mineral assemblages developed in response to burial and heating to similar depths and temperatures. It can represent either the diagnostic mineral assemblages that indicate the physical conditions of metamorphism or the pressure-temperature conditions that produce a particular assemblage in a rock of a specific composition.

Changes in pressure and temperature, due to tectonic activity or intrusion of magma, cause minerals in rocks to react with one another to produce new mineral assemblages and textures. This is the process that transforms igneous and sedimentary rocks into metamorphic rocks. The mineral assemblage that is preserved in a metamorphic rock reflects the composition of the original igneous or sedimentary parent and the pressure-temperature history experienced by the rock during metamorphism. Hence, rocks of the same initial composition that are subjected to the same pressure-temperature conditions will contain identical mineral assemblages.

Facies names and boundaries. The metamorphic facies to which a rock belongs can be identified from the mineral assemblage present in the rock; the pressure and temperature conditions represented by each facies are broadly known from experimental laboratory work on mineral stabilities. Several metamorphic facies names are commonly accepted, and these are shown, along with their approximate pressure-temperature fields, in **Fig.** 4*a*. These facies names are based on the mineral assemblages that develop during metamorphism of a rock with the composition of a basalt, which is a volcanic rock rich in iron and magnesium and with relatively little silica. For example, the dominant mineral of the blueschist facies (in a rock of basaltic composition) is a sodium- and magnesium-bearing silicate called glaucophane, which is dark blue in outcrop and blue or violet when viewed under the microscope. Characteristic minerals of the greenschist facies include chlorite and actinolite, both of which are green in outcrop and under the microscope. Basaltic rocks metamorphosed in the amphibolite facies are largely composed of an amphibole called hornblende. The granulite facies takes its name from a texture rather than a specific mineral: the pyroxenes and plagioclase that are common minerals in granulite facies rocks typically form rounded crystals of similar size that give the rock a granular fabric. *See* AMPHIBOLE; AMPHIBOLITE; BASALT; BLUESCHIST; GLAUCOPHANE; GRANULITE; PYROXENE.

Rocks of different composition show different mineral assemblages in each of the facies. A shale metamorphosed to greenschist facies conditions would contain chlorite, biotite, and perhaps garnet as its key indicator minerals, in contrast to the chlorite + actinolite assemblage that would be found in a rock of basaltic composition. In the amphibolite facies, the indicators in a metamorphosed shale would be garnet, staurolite, or kyanite rather than amphibole. *See* SHALE.

The boundaries separating individual facies from one another are chemical reactions called isograds that transform one key mineral assemblage into another. For example, a typical greenschist is transformed into an amphibolite by the hornblende isograd reaction chlorite + actinolite + plagioclase = hornblende, and an amphibolite can be transformed into a granulite by the reaction hornblende = clinopyroxene + orthopyroxene + plagioclase. In general, these reactions occur over a range of pressure-temperature conditions that depend upon the specific compositions of the minerals involved. In most cases, facies are separated by several reactions that occur within a restricted interval of pressure-temperature space. Thus the boundaries between facies are diffuse and are represented by tinted areas rather than individual lines in Fig. 4. *See* HORNBLENDE; PHASE EQUILIBRIUM.

Facies series. Different geological terranes show different combinations of metamorphic facies. For example, one area may show a progression from blueschist to eclogite to greenschist facies over a distance of a few kilometers (1 km = 0.6 mi), whereas another area may be characterized by a sequence from greenschist to amphibolite to granulite facies. Because the specific facies are related to the pressure and temperature conditions of metamorphism, these different facies series indicate differences in the pressure-temperature histories of the two areas.

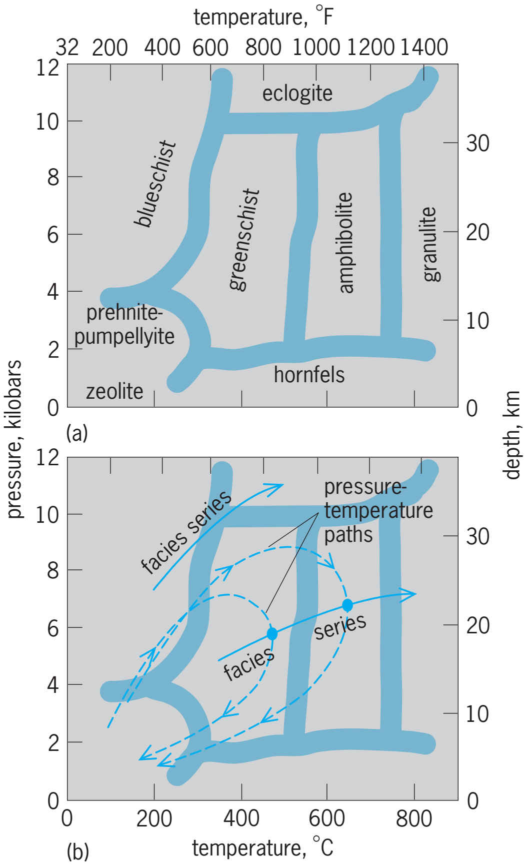


Fig. 4. Pressure–temperature fields. (*a*) Diagram showing the relationship between the different metamorphic facies and the physical conditions of metamorphism. Boundaries (tinted areas) between facies are regions where isograd reactions occur. (*b*) Diagram showing the relationships between metamorphic facies, individual rock pressure–temperature paths, and two typical facies series. Note that the facies series connects maximum temperature points on pressure–temperature paths followed by different rocks. 1 kilobar = 10^8 pascals. 1 km = 0.6 mi.

The blueschist-eclogite-greenschist facies series develops in areas that have experienced high-pressure metamorphism at relatively low temperatures. This type of pressure–temperature history is produced in subduction zones where two lithospheric plates collide with one another. In contrast, the greenschist-amphibolite-granulite series forms in regions that have undergone heating at middle to lower crustal depths. This facies series is exposed in many of the mountain belts of the world (Appalachians, Alps, Himalayas) and appears to result from tectonic thrusting of one continent over another. Lowpressure hornfels or greenschist-amphibolite facies series develop in response to intrusion of hot magma at shallow levels of the Earth's crust; this is a process referred to as contact metamorphism. As illustrated by these examples, metamorphic facies series can be used as an indicator of the plate tectonic history of a geologic region. *See* ECLOGITE; MAGMA; PLATE TECTONICS.

Pressure–temperature paths. Because burial and uplift rates may differ from rates of heat transfer, individual rocks may follow paths through pressure–temperature space that are quite distinct from the facies series recorded by the host terrane. Mineral reaction rates are strongly dependent upon temperature and proceed faster during heating than during cooling; pressure has little effect on these rates. Hence, metamorphic facies and facies series generally reflect equilibration at the thermal maximum of metamorphism, regardless of the complexity of the actual pressure–temperature path followed by the rock in response to tectonic movements. As shown in Fig. 4*b*, individual pressure–temperature paths in many cases intersect the facies series path at a high angle. Disequilibrium mineral assemblages and chemical zoning in some minerals can sometimes be used to reconstruct portions of these pressure–temperature paths and hence gain information on how the depth of burial and the temperature of an individual rock varied through time. This information can be used to refine tectonic interpretations of an area based on the metamorphic facies series. *See* METAMORPHIC ROCKS; METAMORPHISM; PETROLOGY.

Jane Selverstone

Bibliography. A. D. Miall, *Principles of Sedimentary Basin Analysis*, 3d ed., Springer-Verlag, 1999; A. Miyashiro, *Metamorphic Petrology*, 1994; G. Nichols, *Sedimentology and Stratigraphy*, Blackwell, 1999; F. J. Turner, *Meatamorphic Petrology: Mineralogical, Field, and Tectonic Aspects*, 2d ed., 1981; R. G. Walker and N. P. James (eds.), *Facies Models: Response to Sea Level Change*, Geological Association of Canada, 1992.

Facilities planning and design

The design and planning of the physical environment of an activity to best support the execution of this activity. The activity may be one of many different types, such as manufacturing (plant layout), health care (hospital layout), logistics (warehouse design), transportation (airport layout), and services (bank layout, stadium layout).

Facilities design is a typical engineering design project and as such exhibits the main characteristics of any (large) long-term engineering design project. It is a complex problem that is ill defined. Many of the objectives and constraints are implicitly assumed or can only be stated qualitatively. Different shareholders in the design and the use of the facility have a variety of definitions and objectives of the project. Hence, an analytical problem statement or a mathematical solution method, such as linear programming, to find the "best" design does not exist. Many variant and competent designs may be generated, and the selection of the final design will involve both quantitative and qualitative measures. *See* ENGINEERING DESIGN.

Facilities design is an interdisciplinary team effort; no single engineer or engineering discipline can complete the full design or has all the required expertise. Industrial engineers typically are responsible for the logical design of the facility and focus on space allocation and functional flow in the facility, while civil engineers focus on the physical construction. Communication between all the user groups and the design team is essential for the final acceptance of the facility. *See* CIVIL ENGINEERING; INDUSTRIAL ENGINEERING.

Most facilities exist for extended periods of time, which may span many decades. Activities and the requirements for their supporting facilities will change constantly over time, and the function and the detailed layout of the facility will be redesigned many times during its life cycle. A primary characteristic of a good facility design is that it is flexible and easily adaptable. Integration in a spatial and temporal master plan is essential for the future efficient use of the facility.

Projects. Facilities design projects must identify the following two basic elements during the early phases: activity centers and the affinities between these centers in the facility.

Activity center. An activity center, commonly called a department, is a compact area that has a homogeneous material flow with the rest of the facility so that it can be treated as a single unit (with respect to the material flows in the facility). Its area should not be too small, since it is not important enough to be considered separately. A rule of thumb is to have an area larger than 3% of the total facility area. This means that the number of departments should be limited to approximately 35. Its area should also not be too large, since the material flow might no longer be homogeneous over the area of this department.

Affinity. The pairwise relationship between two departments expresses the affinity between these two departments, based on material flow and environmental considerations. Relationships can be numerical (quantitative) if accurate information is available or symbolic (qualitative). The closer one would like these departments to be, the more positive their relationship. A negative relationship means that it is desirable to keep the two departments separate, such

TABLE 1. Qualitative relationships representation

Closeness	Letter	Lines	Color
Absolute necessary	A	4	Red
Especially important	E	3	Orange
Important	I	2	Blue
Ordinary closeness	O	1	Green
Unimportant	U	No	None
Undesirable	X	Wave	Brown

TABLE 2. Qualitative relationship chart

	A	B	C	D	E	F	G
A		E	O	I	O	U	U
B			U	E	I	I	U
C				U	U	O	U
D					I	U	U
E						A	I
F							E
G							

as for noise or vibration pollution. The table of relationships is called a relationship chart or relationship matrix. Possible representations of qualitative relationships are given in **Table 1**. An example of a qualitative relationship chart, as originally proposed by R. Muther, is shown in **Table 2**.

Design sequence. The following sequence of steps can be used to identify departments and their affinities in the project.

1. Identify the major materials-handling flows in the project. This material flow can be of many different types depending on the facilities design application. In a hospital, patients, medical personnel, drugs, and medical supplies are all important material flows. In an airport, passengers (arriving and departing), crew, planes, and luggage are all important material flows. For a bank customer, personnel, documents, and money are examples of important material flows. In manufacturing, parts, personnel, and tools can be examples of the major materials-handling flows. *See* MATERIALS HANDLING.

2. For each of the materials-handling flows identified in step 1, identify the major processing steps. Again, this includes a wide variety of tasks depending on the application. Examples are a customer dining in a restaurant, luggage dispersal to travelers in an airport, drug storage in the pharmacy of a hospital, order picking in a warehouse, and so on. The processing steps are executed in activity centers or departments. At this particular stage, a functional description of the steps is required, such as store, move, or transform.

3. Construct a graph with a column for each of the materials-handling flows and a row for each of the activity centers. For each of the materials-handling flow, a line is drawn between the activity centers to indicate the order in which the centers are visited. An example of such a graph is a multiproduct process chart. *See* METHODS ENGINEERING.

4. Given the materials-handling flows and the activity centers, determine the affinities between the activity centers in a common unit. If only materials-handling flows are present, then this is a relatively easy step. The materials-handling load, such as a pallet of forklift truck trip, might be the common materials-handling unit. If other affinities are also present, then their size must be carefully determined to be consistent with the materials-handling flow units. Examples of positive affinities are the desirability of having windows in offices and cafeterias, which implies a location on the perimeter of the layout. Examples of negative affinities are welding (which produces sparks) and painting (which produces combustible vapors) departments, vibration pollution between a stamping and a measuring department, and noise pollution between a heavy machining and an office department. Once a common handling unit has been defined, the flows in the above graph can be quantified. All materials-handling flows and other affinities are then summarized in a two-dimensional relationship matrix, also called a from–to matrix, which has as elements the sum of the affinities between two departments for all major materials-handling flows.

5. Given the estimated materials-handling flows that need to be processed in each of the activity centers, compute the number of individual servers required in each activity center. Servers are the generic term for the people or machines that execute the processing step in the activity center. In a manufacturing context, a server is typically a machine and all its surrounding support equipment. In a hospital emergency room, a server may be an examination area surrounded by a privacy curtain. In a grocery store checkout area, a server may be the cashier and the station. Deterministic capacity calculation, queuing theory, and digital simulation are three methods used to compute the number of servers, which are ranked by increasing level of accuracy and required effort. Finally, based on the number of servers and the required area for each server and possible space for waiting area for the server, compute the required area for each activity center.

Space-relationship diagrams. Based on the data for the activity centers, such as required area and other constraints, and on the affinities between the centers, a logical diagram of the centers and the material flows can now be constructed. R. Muther defined such a diagram and called it a space-relationship diagram. The objective of the logical design of the facility is to place activity centers with large positive affinities close to each other and to separate centers with significant negative affinity, subject to space and site constraints. Three major types of diagrams can be identified. For historical reasons their names are based on their implementation in manufacturing activities, but the prototypes apply to all types of facilities.

Product layout. A product layout is recommended when most of the materials or products follow the same material flow path. This corresponds to a flow shop in manufacturing. Most warehouses also correspond to a flow shop, since the product inventory has a limited number of paths from receiving to shipping. The objective is to have a simple configuration for the major material flow path, such as straight line or L- and U-shaped. The activity centers are located according to the sequence of processing steps required by the product.

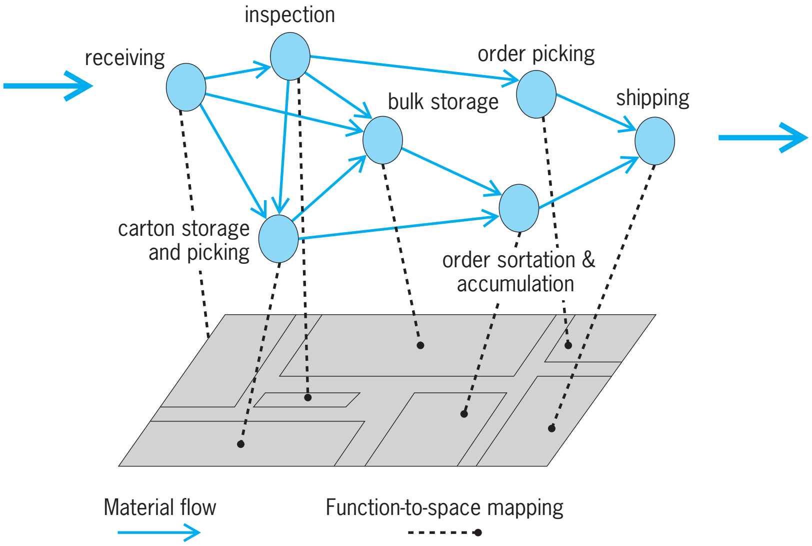
Process layout. A process layout is recommended when products have different paths through the facility. This corresponds to a job shop in manufacturing. The typical example is a repair shop with different types of machining centers. The objective is to minimize the total amount of material movement. Because of the complexity of the combined flow paths, optimization models can provide some support in the early design phases. Since accurately predicting the future material flows over the lifetime of the facility is nearly impossible, the facility should be designed to accommodate reconfiguration.

Project layout. A project layout is recommended when materials are transported to a site to be integrated into a final product. A typical example is ship construction.

Conceptual layout. In the next phase, the space-relationship diagram is converted in a conceptual layout, where departments are shown without internal details and without material flow corridors. This conversion is most often executed by the design engineer with drawing computer-aided design (CAD) tools but without formal problem definition or algorithmic support. In recent years, significant progress has been made by the academic community in formalizing and automating this conversion. But these advances have not been transferred to the practice of facilities design. An example of mapping from the functional flow diagram to the conceptual layout is shown in the **illustration**. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Detailed layout. The conceptual layout is then converted into a final detailed layout. Space is added for materials-handling corridors and for support functions, such as restrooms, utility closets, and so on. The layout elements corresponding to the requirements of federal regulations such as the Occupational Safety and Health Administration (OSHA), Environmental Protection Administration (EPA), Americans with Disabilities Act (ADA), and of local fire and safety regulations are integrated into the design. If structural elements such as support pillars are present, the design is adapted to incorporate these constraints.

Evaluation. During the design process, numerous engineering design decisions are made that may result in many alternative layouts. The comparative evaluation of facility designs is based on a mixture of engineering economy and qualitative factors. All quantifiable costs for a particular design over the expected life span of the facility are computed and plotted along a time line. This includes one-time construction and material transportation equipment purchasing costs, as well as recurring operating costs. Typical examples of recurring costs are those for energy and material transportation. Costs occurring in the future are reduced to their net present value (NPV) based on a monetary discount factor. Since a facility supports a particular activity, the net present value of the initial investments and recurring costs typically are negative. The facility with the least negative NPV, subject to budget constraints, is preferred on pure economic grounds. However, a number of qualitative evaluation criteria are typically present. A weight is assigned to the economic evaluations and the qualitative evaluations and the facility design with the highest weighted score is finally selected. The weights are based on a polling of the different stakeholders in the facility and are not uniquely defined or constant over time. To provide support to the decision makers, sensitivity analysis on the weight has to be performed.



Mapping from the functional flow diagram to the conceptual layout. (*After L. F. McGinnis et al., 2000*)

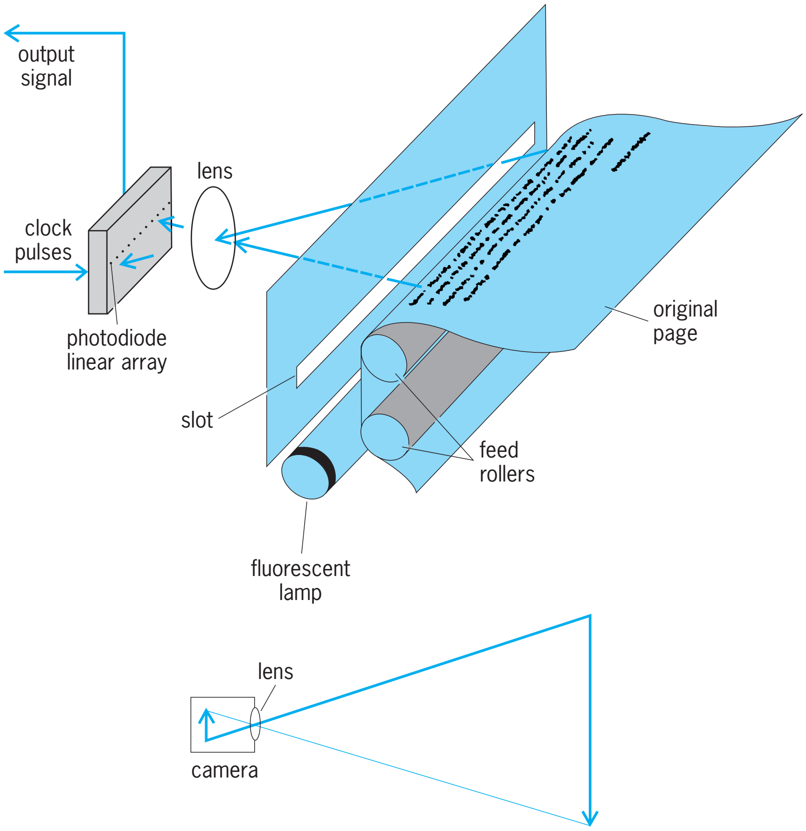
Simulation tools. It is clear that the above facilities design method incorporates many ill-defined objectives and constraints. Consequently, there has been a very limited development of computer-based facilities design tools beyond drawing tools and facilities management tools. The most often used software during the design of facilities is digital simulation coupled with animation. Simulation is capable of very detailed analysis of a proposed facility design and of the evaluation of some minor design variations, but it cannot generate significantly different designs. As a result, the design of facilities remains largely a person-based activity supported by computer-based drawing and simulation tools. *See* COMPUTER-AIDED ENGINEERING; SIMULATION.

Marc Goetschalckx

Bibliography. R. L. Francis, L. F. McGinnis, and J. A. White, *Facility Layout and Location: An Analytical Approach*, 2d ed., Prentice Hall, 1992; S. Heragu, *Facilities Design*, PWS Publishing, 1996; J.-G. Kim and M. Goetschalckx, An integrated approach for the concurrent determination of the block layout and i/o point locations based on the contour distance, *Int. J. Prod. Res.*, 43(10):2027–2047, 2003; S. Konz, *Facilities Design*, Wiley, 1985; A. Kusiak and S. S. Heragu, The facility layout problem, *Eur. J. Oper. Res.*, 29:229–251, 1987; L. F. McGinnis et al., Rethinking warehouse design research, *Proceedings of the Material Handling Research Colloquium*, York, June 10–14, York, PA (published in R. Graves et al., eds., *Progress in Material Handling Research: 2000*, pp. 206–216, 2000; R. Meller and K.-Y. Gau, The facility layout problem: recent and emerging trends and perspectives, *J. Manuf. Sys.*, 15(5):351–366, 1996; H. Sherali, B. M. P. Fraticelli, and R. D. Meller, Enhanced model formulations for optimal facility layout, *Oper. Res.*, 51(4):629–644, 2003; R. Muther, *Systematic Layout Planning*, 2d ed., Cahners Books, 1973; D. R. Sule, *Manufacturing Facilities: Location, Planning and Design*, 2d ed., PWS Publishing, 1988; J. A. Tompkins et al., *Facilities Planning*, 3d ed., Wiley, 2003.

Facsimile

The process by which a document is scanned and converted into electrical signals which are transmitted over a communications channel and recorded on a printed page or displayed on a computer screen. The scanner may be compared with a camcorder, and the recorder is similar to an office copier or a computer printer. As an alternative to scanning, a document stored in computer memory can be transmitted. As an alternative to recording, a text facsimile (fax) image can be captured in computer memory and converted into computer-processable text by optical character recognition (OCR) software. Telephone lines or satellites provide the communication channel. More than 99% of facsimile units are International Telegraph and Telephone Consultative Committee Group 3, used for sending business documents. The remainder have specialized designs tailored for a number of important applications that cannot be fulfilled by Group 3.



Photodiode facsimile scanner.

Group 3 facsimile. Group 3 digital facsimile handles far more messages than any other communications system except voice. Most facsimile units communicate over the Public Switched Telephone Network, alternatively called the General Switched Telephone Network. A built-in high-speed digital modem automatically selects the highest modem speed (28,800–2400 bits/s) common to both facsimile units. If the telephone-line quality is not good enough for this transmission speed, a lower speed is negotiated during initialization. A very small percentage of the fax units send or receive via a cellular phone connection. Group 3 standards support alternative use of the Integrated Services Digital Network, including the 64-kilobit/s channel. *See* INTEGRATED SERVICES DIGITAL NETWORK (ISDN); MODEM; TELEPHONE SERVICE.

Unattended-reception capability allows transmission with automated recording. Some fax units can place calls and send documents without an operator. The cost of a facsimile call usually is far less than a voice call since facsimile sends information much faster, at a rate of 3–10 or more pages per minute.

Scanning. An image of the original page is formed by a lens in a way similar to that of an ordinary camera (see **illus.**). A charge-coupled device linear array of small photodiodes is substituted in the facsimile scanner for the camera film. The portion of the image falling on the linear diode array is a thin line, 0.005 in. (0.13 mm) high, across the top of the page being transmitted. Typically, 1728 diodes are used to view this line for a page $8^{1}/_{2}$ in. (216 mm) wide. The photodiode corresponding to the left edge of the page is first checked to determine whether the very small portion of the image it detects is white (the paper background) or black (a mark). The spot detected by a single photodiode is called a picture element (a pel for short if it is recorded as either black or white, or a pixel if a gray scale is used). Each of the 1728 diodes is checked in sequence, to read across the page. Then the original page is stepped the height of this thin line, and the next line is read. The step-and-read process repeats until the whole page has been scanned. *See* CAMERA; CHARGE-COUPLED DEVICES; PHOTODIODE; TELEVISION CAMERA.

Another class of flatbed scanner uses a contact image sensor linear array of photodiodes whose width is the same as the scanned width. One version has a linear array of fiber-optics rod lenses between the page being scanned and the sensor array. Light from a fluorescent lamp or a linear light-emitting-diode array illuminates the document beneath the rod lenses. The reflected light picked up by the sensor generates a signal that is proportional to the brightness of the spot being scanned. A second version has a hole in the center of each square pixel sensor element. Light from a light-emitting diode passes through this hole to illuminate the area of the document page at this pixel. No lenses or other optical parts are used. *See* FIBER-OPTICS IMAGING; LIGHT CURVES.

In drum-type scanning, the original sheet of paper is mounted on a drum that rotates while the scan head with a photosensor moves sideways the width of one scanning line for each turn of the drum. Drum-type scanners are used mainly for remote publishing facsimiles and for color scanning in graphic arts systems.

Recording. In the recording process, facsimile signals are converted into a copy of the original. Facsimile receivers commonly print pages as they are received, but in an alternative arrangement pages may be stored and viewed on a computer screen. Facsimile capability is possible with computers, including portable varieties.

One of the most common methods is thermal recording. There are 1728 very fine wires positioned in a row across the recording paper. These wires touch the paper and produce very small hot spots as current passes through them. These hot-spot sections of the wires form a straight line across the page at a resolution of 200 dots per inch (8 dots per millimeter). A wire takes only a few milliseconds to go from a cool, nonmarking condition to a hot, marking condition. Each recording wire corresponds to a photodiode in the linear array of the transmitting scanner. Direct-writing, thermally sensitive coated paper is used to produce the image, but a more expensive system of thermal transfer that uses ordinary white paper is offered in some units.

Xerographic recording with a laser may be used where a higher resolution of 400 dots per inch and higher writing speeds are needed. The laser beam, modulated by the picture signal, is swept in a line across the photoreceptor drum (similar to the operation of laser printers used with computers). The drum steps ahead one recording line height of 0.0025 in. (0.064 mm) before writing the next line. The toning, fixing, and paper-feed systems resemble those of computer printers. A linear array of light-emitting diodes or liquid-crystal shutters may be used instead of the laser beam. *See* COMPUTER PERIPHERAL DEVICES; LASER; PHOTOCOPYING PROCESSES.

Ink-jet recording is one of the least expensive methods of plain-paper recording. An ink-jet facsimile recorder sprays ink droplets on plain paper to reproduce black markings scanned at the facsimile transmitter. At resolutions of up to 300 dots per inch, the copy quality compares favorably with that produced on a laser printer. The ability to print either inch-based resolution or metric resolution on the same printer gives the ink-jet printer an advantage.

Advanced Group 3 capabilities. Group 3 standards can support computer functions with fax, adding many optional features without impairing compatibility with Group 3 fax machines that lack them.

Group 3C. This 64 kb/s option for Group 3 facsimile on public digital networks uses full-duplex or half-duplex communication with the circuit-switched mode and error-correction mode. This method is considerably simplified compared with the protocols used by Group 4 facsimile. Sending or receiving documents from another G3-64 fax machine is at 64 kb/s using an Integrated Services Digital Network B channel. Communication with a standard Group 3 machine on the Public Switched Telephone Network is possible via the Integrated Services Digital Network.

Cellular radio. Automobiles equipped with cellular radio enable communication by facsimile as well as by telephone. The brief loss of signal that occurs when switching between cells is heard as a click and the voice signals continue normally, but with facsimile each click causes a streak across the page and may obliterate information. The frequency of cell changing is unpredictable, and facsimile pages may have no streaks or up to 10 streaks per page. Switching may also result in a failure message at the facsimile transmitter, caused by too many received errors, or may cause a disconnect. Cellular networks also are subject to loss of signal, multipath, and fading due to physical obstructions, reflections from buildings, and distance. Use of the error-correction mode is desirable. *See* MOBILE RADIO.

Computer facsimile. Personal computers can emulate Group 3 facsimile for transmission of text and graphic files. A so-called fax board is plugged into a vacant slot in the computer, and a software program converts computer language files into Group 3 facsimile files. A standard facsimile modem on the fax board makes the signals and the communication protocol identical with Group 3. Documents sent to or received from Group 3 facsimile or other personal computers can be viewed and edited on the display screen, or a hard copy can be made on the computer's printer. By adding a scanner, documents can also be sent (for complete emulation of Group 3 facsimile). *See* MICROCOMPUTER.

Computer file transfer. Mainframes, file servers on local area networks, workstations, business computers, personal computers, and even palm-size personal digital assistants often have the ability to send and receive Group 3 facsimile documents. Several options additionally allow automatic transmission of data by the Group 3 facsimile high-speed modem and worldwide standardized facsimile protocols. High-level data-link control (HDLC) frames are sent as the logical equivalent of an error-corrected facsimile message, with the protocols and transmission speeds being automatically established. This file transfer method sends any data file with or without additional information concerning the file.

Other facsimile systems. A number of facsimile systems besides Group 3 are in operation.

Group 4 facsimile. This system sends error-free copies over public data networks at 200 and 300 lines per inch (7.9 and 11.8 lines per millimeter), with 240 and 400 (9.4 and 15.7) optional. Group 4 operates over high-speed error-free digital networks using the Open Systems Interface (OSI) model. Communication is also possible with Group 3 units by a so-called dual-mode Group 4 Class 1 fax. Dual-mode Group 4 facsimile may have standard Group 3 with a fax modem, Group 3c (64-kb/s capability), or both. There are slight page size distortions caused when interworking between Group 3 metric-based resolutions and similar Group 4 inch-based resolutions. Group 3 now provides most of the same features as Group 4b, using simplified protocols.

Newsphoto facsimile. Most pictures printed in newspapers have been sent by facsimile. News services have nationwide networks for distribution of newsphotos. Thousands of newspapers receive the latest news pictures at the same time by a broadcast operation, coming from almost anywhere in the world. These systems reproduce the gray scale very well, with each pixel representing one of 64 shades. Color separations are also sent as a set of four gray-scale originals representing the primary colors plus black. An electronic picture desk may store the pictures in memory as they are received.

Remote publishing. Newspapers and magazines are published from full-page facsimile masters received in plants thousands of miles away from the place where the pages are composed. The facsimile recording is made on a film negative or directly on a printing plate. High-resolution, high-speed facsimile units send full-size newspaper pages over wideband satellite or land lines. Resolutions of 800–1000 lines/inch may be used. Major newspapers transmit to many printing plants via direct satellite broadcast to receiving stations located at each plant. Color pages are sent by making three or four transmissions of the original color page by using color separation filters.

Weather facsimile. Weather networks cover the United States, constantly making copies of weather charts. Cloud-cover photographs and multispectral band images are recorded by facsimile from signals sent by orbiting and geostationary satellites. *See* METEOROLOGICAL SATELLITES. Kenneth R. McConnell

Bibliography. K. R. McConnell, *Facsimile Technology and Applications Handbook*, 2d ed., 1992.

Factor analysis

A method of quantitative multivariate analysis with the goal of representing the interrelationships among a set of continuously measured variables (usually represented by their intercorrelations) by a number of underlying, linearly independent reference variables called factors. Although the term factor analysis has come to represent a family of analysis methods, the two most commonly used approaches are the full component model, in which the entire variance of the variables (represented by unities inserted in the principal diagonal of the correlation matrix) is analyzed, and the common factor model, in which the proportion of the variance that is accounted for by the common factors (represented by communality estimates inserted in the principal diagonal) is analyzed.

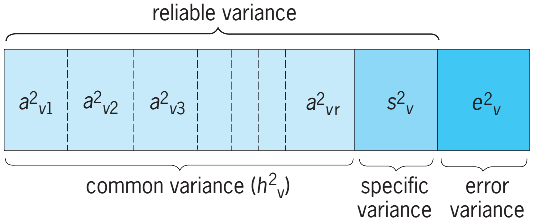


Fig. 1. Diagram of the total variance of a variable in standard score units subdivided into types of variances.

The method was developed in England around the turn of the century and was first applied to the study of the structure of intellectual abilities. Since then it has been used in many disciplines, from agriculture to zoology, in which the underlying structure of multiple variables and their representation in that structure are of interest. Another application of factor analysis is to represent parsimoniously the variables in the set on which the observations are made by a smaller number of underlying reference variables or factors.

Equations and matrices. The defining equation for the full component model is shown in Eq. (1), where

$$z_{iv} = a_{v1}F_{1i} + a_{v2}F_{2i} + \cdots + a_{vr}F_{ri} \qquad (1)$$

z_{iv} represents the standard score for case i on variable v, a_{vr} is the weight for variable v on factor r, and F_{ri} is the amount (factor score), in standard scores, on factor r possessed by case i.

The defining equation of the common factor model is shown in Eq. (2), where the first elements

$$z_{vi} = a_{v1}F_{1i} + \cdots + a_{vr}F_{ri} + a_{vu}U_{iv} \qquad (2)$$

are the same as in the full component model, a_{vu} is the weight given variable v's unique factor, and U_{iv} is case i's unique factor score for variable v. Each variable's unique factor includes random errors, influences affecting that variable only, and all other sources of error and bias that do not contribute to the correlations of that variable with other variables in the set, and therefore do not enter into determining common factors. Usually, only the common factors are evaluated from the data.

The total variance of each variable, $\sigma^2{}_v$, may be separated into its constituent variances, as represented by Eq. (3).

$$\sigma_v^2 = \sigma_{v1}^2 + \sigma_{v2}^2 + \cdots + \sigma_{vr}^2 + \sigma_{vs}^2 + \sigma_{ve}^2 \qquad (3)$$

Dividing both sides of the equation by $\sigma^2{}_v$ to convert to standard scores yields Eq. (4) in terms of

$$1.0 = a_{v1}^2 + a_{v2}^2 + \cdots + a_{vr}^2 + s_v^2 + e_v^2 \qquad (4)$$

proportions of variance. This set of proportions of variances may also be represented diagrammatically (**Fig. 1**).

In applying the common factor model, usually only the common variances are evaluated, and their sum for a variable is represented by $h^2{}_v$ and is termed the communality of the variable. The full component model accounts for the components making up the total variance. In practice, however, a "truncated" analysis is usually performed, so that only the common variances (or factors) are extracted from the correlation matrix, even though the components model is used.

The correlation r_{jk} between two variables can be accounted for in terms of their common factor loadings. For orthogonal factors (axes) this can be represented by Eq. (5).

$$r_{jk} = a_{j1}a_{k1} + a_{k1} + \cdots + a_{jr}a_{kr} \qquad (5)$$

Thus the correlation between two variables can be accounted for in terms of their common factors. The corresponding equation in matrix notation for a set of variables is $R = FF'$.

The correlation between two variables may also be represented geometrically by the scalar product of the length of two vectors and the cosine of the angle (ϕ) between them. This relationship is represented by Eq. (6). If each variable is indicated by a

$$r_{jk} = V_j V_k \cos \phi_{jk} \qquad (6)$$

unit length vector (representing its total variance), the correlation is equal to the cosine of the angle between them. Representations of several r's in terms of unit length vectors and the angle between them are shown in **Fig. 2**. Since the cosine of an angle of 90° is 0.0, an r of zero may be represented by two unit length vectors separated by an angle of 90°, and so on. Positive correlations have angles between 0 and 90°, negative correlations between 90 and 180°.

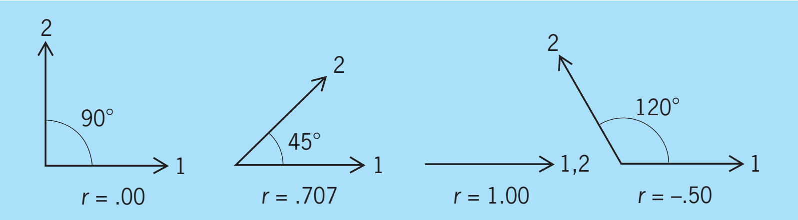


Fig. 2. Vectorial representation of correlation coefficients.

Figure 3 depicts an attempt to represent the following correlation matrix geometrically in two dimensions:

Variable	1	2	3
1	1.000	.000	.707
2	.000	1.000	−.505
3	.707	−.500	1.000

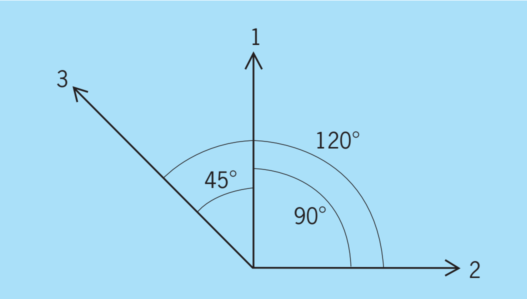


Fig. 3. Representation demonstrating that the three-variable correlation matrix cannot be represented correctly in two dimensions.

Since the sum of 45° plus 90° is not equal to 120°, the representation is erroneous; it requires three dimensions to represent the vectors and the angles between them correctly. The minimum number of dimensions required to represent a matrix of intercorrelations may be taken to indicate the rank of the matrix. One of the primary uses of factor analysis is to determine the number of independent reference variables or dimensions needed to account for the intercorrelations of a set of continuous variables. With empirical data that contain measurement and other sources of error, the rank can only be approximated.

Figure 4 is a geometric representation of the following matrix of intercorrelations among four variables, which can be represented correctly in two dimensions:

Variable	1	2	3	4
1	(1.00)	.80	.96	.60
2	(.80)	(1.00)	.60	.00
3	(.96)	.60	(1.00)	.80
4	(.60)	.00	(.80)	(1.00)

A pair of orthogonal cartesian axes has been inserted into the vector space, with axis I coinciding with vector 2, and axis II coinciding with variable 4.

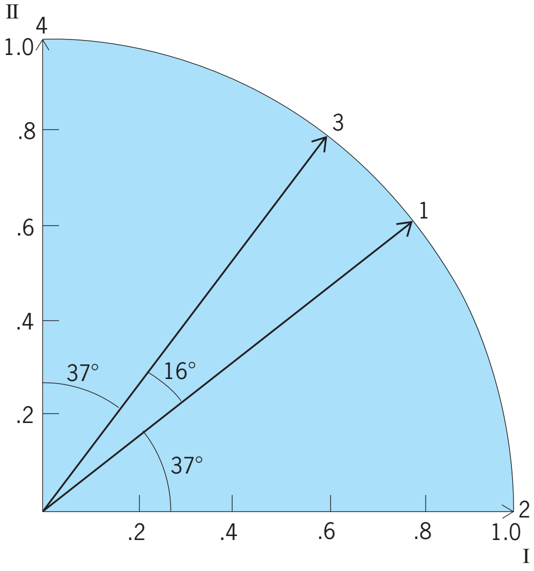


Fig. 4. Vectorial representation of the intercorrelations among four variables in a two-dimensional space.

If it is preferred to have the reference axes (which actually extend from +1.0 to −1.0) in another position relative to the variable vectors (on either orthogonal or oblique axes), they may be transformed to that position by "rotating" them about the origin. When an analysis has been performed, the resulting values are usually written in an abbreviated factor matrix form, with the entries in the body of the table (that is, the factor coefficients) being the coordinates of the variables on the axes, and h^2 (the sum of the squared factor coefficients on orthogonal axes) representing the communality of the variable:

	Factor		
Variable	I	II	h^2
1	.8	.6	1.0
2	1.0	.0	1.0
3	.6	.8	1.0
4	.0	1.0	1.0

For fallible data the values of the communalities would be less than 1.0. The nature of the underlying characteristic represented by a factor is frequently inferred from the common element or characteristic of the variables having high loadings on the factor.

Procedure. The most commonly performed steps in factor analysis are as follows:

1. *Factoring the correlation matrix*: The most commonly used procedure is to obtain the eigenvectors and eigenvalues of the matrix by least-squares procedures with unities inserted in the principal diagonal (principal components), or with communality estimates in the principal diagonal (principal factors). In the latter case the factoring process is frequently iterated one or more times (using the communalities computed from the previous factoring) to stabilize the communality estimates. A number of other factoring procedures have been devised, including diagonal, centroid, multiple-group, image, alpha, nonlinear, nonmetric, and maximum-likelihood solutions.

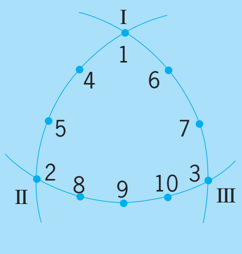
2. *Determining the number of factors*: Since, with falliable data, there is not an exact number of factors, a large number of approximate criteria have been developed. Some commonly used ones with principal-components analysis are retaining those factors with eigenvalues ≥ 1.0, and Cattell's screen test.

3. *Rotation and interpretation of factors*: The position of the axes determined by the mathematical factoring methods are usually not optimal for inferring the nature of the factors based on the content of the variables and knowledge of the field of investigation. Axes may be rotated about the origin, yielding orthogonal or oblique axes solutions according to several models that are more appropriate for interpretation.

Historically, one of the first models was C. Spearman's two-factor theory. He observed that measures of mental abilities intercorrelated positively, and developed a theory of a general intellectual ability, which he called the two-factor theory. All the mental ability measures loaded on a common factor (labeled "g," for general intellectual ability), and each test also had its own specific factor. K. J. Holzinger modified the model to the bifactor theory. In addition to the general or g-factor, groups of the mental ability tests have factors in common; these group factors represented cognitive abilities such as verbal, numerical, and reasoning.

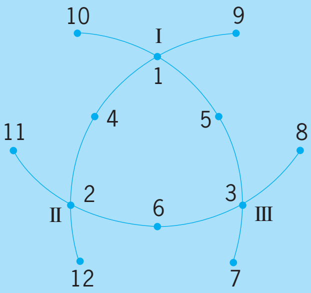
A more general model is L. L. Thurstone's multiple-factor theory, which is based on the principle of "simple structure." For the case where each variable is factorially less complex than the set of variables taken together, the following are some considerations for evaluating a simple structure: Each variable should have at least one factor loading close to zero. Each factor should have at least as many near-zero loadings as there are factors. For every pair of factors there should be several variables with near-zero loadings on one factor, but not on the other. For analyses with four or more factors, a large proportion of the variables should have negligible loadings on any pair of factors. Only a small number of variables should have appreciably large loadings on any pair of factors.

A secondary consideration in guiding rotations of factors based on sets of variables with positive intercorrelations is "positive manifold"; that is, after rotation of axes all factor loadings are zero or positive. **Figure 5** shows a possible simple structure for 10 variables with a positive manifold configuration in three factors on the surface of a sphere, and **Fig. 6** shows one for 12 variables, with bipolar factors having negative as well as positive loadings. Several computer programs have been devised for axis rotation to simple structure. The most widely used one for orthogonal solutions is H. F. Kaiser's normalized-varimax algorithm, and for oblique rotations the oblimin, biquartimin, promax, and maxplane methods. In an oblique solution, the degree of obliqueness of the axes, as well as the matrix of factor coefficients, must be taken into consideration. The matrix representing the obliquities of the axes



	factor pattern		
variable	I	II	III
1	+		
2		+	
3			+
4	+	+	
5	+	+	
6	+		+
7	+		+
8		+	+
9		+	+
10		+	+

Fig. 5. A simple structure configuration for 10 variables in three dimensions, shown on the surface of a sphere, with reference axes I, II, and III at the intersections, and the corresponding factor-pattern matrix.



	factor pattern		
variable	I	II	III
1	+		
2		+	
3			+
4	+	+	
5	+		+
6		+	+
7	–		+
8		–	+
9	+	–	
10	+		–
11		+	–
12	–	+	

Fig. 6. A simple structure for 12 variables in three dimensions with bipolar factors, and the corresponding factor-pattern matrix.

in an oblique solution may be factored to obtain a second-order factor solution, and if several oblique factors are obtained, higher-order analyses may be performed.

"Procrustes" is another approach to rotation of axes. It involves setting up a hypothesized factor pattern for a set of variables based on theory or previous results, and rotating the factor matrix as closely to the hypothesized matrix as possible while still reproducing the correlation matrix as accurately as can be done from the unrotated factors. This procedure is also used to make factor solutions based on common sets of variables as congruent as possible before relating factors across studies.

Factor scores. It will be recalled that the defining equations given at the beginning of this article included F_{ri}'s, or factor scores in standard score units. Since the standard scores on the variables and the factor coefficients are available after a factor analysis, the weights for factor scores can be solved for, or estimated, by linear regression. Theoretically the factor scores can be solved for exactly in the full components model, and only estimated in the common factor model, since estimates of the unique factor scores and weights are usually not available in that model. A more direct, less formal approach to estimating factor scores is to use one or more of the variables that have high coefficients on a factor, but with only negligible coefficients on other factors, to represent the factor.

Statistical inference. The usual statistical approach to factor analysis is descriptive, although some tests of significance and other inferential procedures have been developed. Factor analytical studies may be classified as exploratory or confirmatory. In connection with the latter, analysis of covariance structures is a procedure that is used to test the fit of the data to a model of specified factor coefficients. The number of cases used to obtain the sample data is usually several times the number of variables in order to minimize problems of statistical significance and help ensure that the correlation matrix will be gramian.

Designs. While the intercorrelation of variables over cases on one occasion (R-technique) is the most common design, others such as Q-technique (the correlation of cases over variables on one occasion), and P-technique (the correlation of variables over occasions for one case) have been used. Also, three-mode factor analysis permits the simultaneous analysis of variables, cases, and occasions. Factor analysis has been used in many fields of science to infer the underlying structure of overlapping measured variables, and especially in psychology to analyze measures of individual differences and their underlying structure. *See* STATISTICS. Benjamin Fruchter

Bibliography. D. Child, *The Essentials of Factor Analysis*, 2d ed., 1991; A. L. Comrey and H. B. Lee, *A First Course in Factor Analysis*, 2d ed., 1992; R. P. McDonald, *Factor Analysis and Related Methods*, 1985; R. A. Reyment and K. G. Joreskog, *Applied Factor Analysis in the Natural Sciences*, 2d ed., 1993.

Fagales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the superorder Rosidae of Eudicotyledon. The order consists of 8 families (Betulaceae, Casuarinaceae, Fagaceae, Juglandaceae, Myricaceae, Nothofagaceae, Rhoipteleaceae, Ticodendraceae) and approximately 30 genera and nearly 1000 species. The Fagales are either simple or compound-leaved, woody plants. Flowers are mostly

Butternut or white walnut (*Juglans cinerea*), a North American species showing compound leaves and slender drooping male catkins. (*Ken Sytsma, University of Wisconsin*)

unisexual and much reduced for wind pollination. The female flowers produce one- or two-seeded nut fruits (for example, acorn, chestnut, walnut); the male flowers are grouped into pendant catkins (see **illus.**). Birch (*Betula*), beech (*Fagus*), walnut (*Juglans*), and oak (*Quercus*) are members of the Fagales. *See* BEECH; BIRCH; MAGNOLIOPHYTA; OAK; PLANT KINGDOM; ROSIDAE.

K. J. Sytsma

Falconiformes

A worldwide order of diurnal predacious birds without obvious affinities to other orders of birds. They are not closely related to the owls; any similarities between these two orders are the result of convergence. *See* AVES.

Classification. The Falconiformes are divided into two suborders. The first, Accipitres, has three families: Pandionidae (osprey; one species; worldwide, including Australia; *see* **illustration**); Accipitridae (hawks, eagles, kites, harriers, and Old World vultures; 217 species; worldwide); and Sagittariidae (secretary bird; one species; Africa). The second suborder, Falcone, contains a single family, Falconidae (falcons, caracaras; 62 species; worldwide). The New World vultures, which had been included in the Falconiformes, are now placed with the Ciconiiformes. *See* CICONIIFORMES.

The inclusion of all these families in one order has been questioned. Some scientists conclude that the falcons are not related to the Accipitres but are allied to the owls. The position of the largely ground-dwelling and hunting secretary bird has been questioned, with some researchers suggesting that it may be a predacious member of the Gruiformes. *See* GRUIFORMES; STRIGIFORMES.

Osprey (*Pandion haliaetus*). (*Photo by Glenn and Martha Vargas, © 2002 California Academy of Sciences*)

Characteristics. The falconiforms (or diurnal raptors) range in size from sparrow-sized falconets (*Microhierax*) of southeastern Asia to the gypaetine vultures with wingspans of 8 ft (2.5 m); most are medium-sized to larger birds. They have strong feet, usually with three toes pointing forward and one pointing backward, and ending with sharp claws. Their beaks are hooked and powerful. Their well-developed wings vary in shape according to the type of flight; they are long and pointed in falcons, short and broad in accipitrine hawks, and long and broad in vultures, eagles, and buzzards. Falconiforms generally hunt from the air, with their wing shape depending on the mode of hunting flight, and they feed on animal prey from insects to sizable vertebrates, including fish, and on carrion. However, the palm-nut vulture (*Gypohierax angolensis*) of Africa is partly vegetarian, eating the fruits of oil palms.

Hawks are flying specialists, although most species can walk well. They are monogamous, with a strong pair bond, and they may mate for life. The young are downy, stay in the nest, and are cared for by both parents. Many northern species are migratory, with spectacular concentrations observed along mountain ridges and the coast during the fall migration.

The osprey (see illustration) is specialized for catching fish with a reversible fourth toe, well-developed claws, and spiny scales along its toes. The one species is found almost worldwide.

The secretary bird of Africa is long-legged and specialized for a terrestrial way of life. It runs rapidly, flies only infrequently, and catches poisonous snakes as well as many other vertebrates.

Falcons are fast-flying birds with long, pointed wings. They hunt from the air, diving and catching their prey in the air or on the ground. They possess an extra toothlike projection on the lateral edges of the upper jaw just behind the terminal hook.

Fossil record. Hawks, falcons, and ospreys are known from Eocene-Oligocene times and are well represented in the fossil record. Secretary birds, now restricted to Africa, are known from the Oligocene and Miocene of France. Most interesting is that the Old World vultures (Gypaetinae) have an extensive fossil record in North America.

Economic significance. Hawks had once been regarded as vermin, but now that their value in controlling rodent pests is understood, they are widely protected. The presence of concentrations of dichlorodiphenyltrichloroethane (DDT) and other pesticides in their food has caused drastic decreases in the number of many species and the near elimination of the osprey, peregrine falcon, and bald eagle over large parts of their range. With decrease in the use of DDT and other pesticides, they are regaining their former numbers. Still, many species remain endangered because of habitat loss, decrease in prey, and continued use of toxic pesticides. More recently, several species of Old World vultures in India have decreased greatly in number with the apparent cause being susceptibility to antibiotics present in the dead cattle eaten by these scavengers; the drugs have been heavily overused and had built up in the body of the cattle. *See* PESTICIDE. Walter J. Bock

Bibliography. L. Brown and D. Amadon, *Eagles, Hawks and Falcons of the World*, Country Life Books, Feltham, 1968; T. J. Cade, *The Falcons of the World*, Cornell University Press, 1982; J. Ferguson-Lees et al., *Raptors: An Identification Guide to the Birds of Prey of the World*, Houghton-Mifflin, 2001.

Fallopian tube

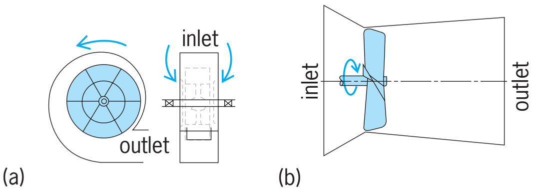
The upper part of the female oviduct present in humans and other higher vertebrates. The fallopian tube extends from the ovary to the uterus and transports ova from the ovary to the cavity of the uterus. Each tube is about 5 in. (12.5 cm) long; one lies on either side of the uterus and is attached at the upper portion. Each curves outward to end in a hoodlike opening, the infundibulum, with many fingerlike projections, the fimbriae; the cavity of the fallopian tube is continuous with the cavity of the coelom. The ovaries lie below and inside the tubal curve. When ovulation occurs, about the middle of each menstrual cycle in humans, the ovum is picked up by the fimbriae and drawn into the infundibulum and hence the oviduct. The lining of this tube is an epithelium with many hairlike processes, cilia, on its surface. These cilia, and probably muscular action, move the ovum along the tube. The ovum remains viable for about 1–3 days only. If fertilization occurs, the ovum moves into the cavity of the uterus and then implants on its wall. If fertilization fails to occur, the ovum degenerates in the uterus. Occasionally, a fertilized ovum fails to enter the uterus, or may be freed into the abdominal cavity, so that an ectopic pregnancy results if the ovum finds a site for implantation. *See* PREGNANCY DISORDERS; REPRODUCTIVE SYSTEM. Walter Bock

Fan

A fan moves gases by producing a low compression ratio, as in ventilation and pneumatic conveying of materials. The increase in density of the gas in passing through a fan is often negligible; the pressure increase or head is usually measured in inches of water.

Blowers are fans that operate where the resistance to gas flow is predominantly downstream of the fan. Exhausters are fans that operate where the flow resistance is mainly upstream of the fan.

Fans are further classified as centrifugal or axial (**Fig. 1**). The housing provides an inlet and an outlet and confines the flow to the region swept out by the rotating impeller. The impeller imparts velocity to the gas, and this velocity changes to a pressure differential under the influence of the housing and ducts connected to inlet and outlet.



(a) (b)

Fig. 1. Fan types. (*a*) Centrifugal. (*b*) Axial.

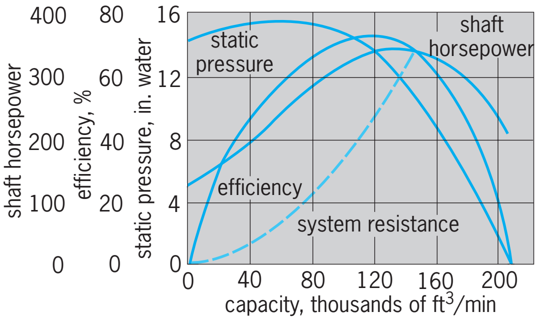


Fig. 2. Static pressure characteristics of backwardly curved blade centrifugal fan at constant speed. 1 horsepower = 746 W. 1 in. water = 249 Pa. 1 ft^3/min = 0.00047 m^3/s.

Performance. In selecting a fan for a particular application, requirements of primary interest are the quantity *Q* of gas to be delivered by the fan and the head *H* which must be developed to overcome the resistance to flow of the quantity *Q* in the connected system. These operating conditions establish the fan dimensions of diameter *D* and rotational speed *N*. Performance of a fan of diameter *D* is rigorously described by its characteristic curves (**Fig. 2**). Fans of different types and sizes are conveniently compared by converting their characteristics to dimensionless form. **Figures 3** and **4** show selected fan performance on the percentage basis, where 100% rating is defined as the peak of the efficiency curve.

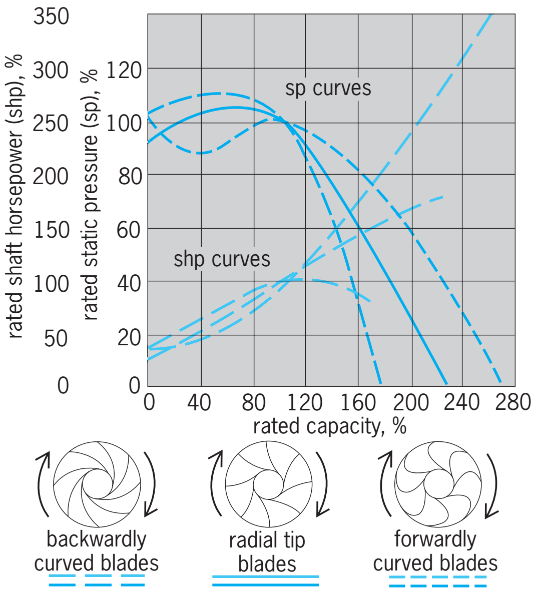


Fig. 3. Percentage characteristics compare performance of three forms of centrifugal fans. 1 horsepower = 746 W.

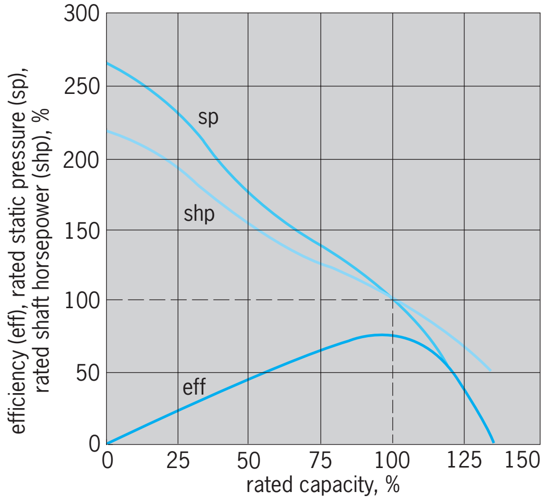


Fig. 4. Percentage characteristics of one axial fan. 1 horsepower = 746 W.

The load placed on the fan must correspond to a condition on its operating characteristic. If the system in which the fan operates presents less resistance than is overcome by the head developed by the fan at the required capacity, more resistance must be introduced by a damper, or the fan speed must be changed, or the excess capacity must be diverted elsewhere.

Performance of a given fan, operating at a given point on its efficiency curve, varies with speed in accordance with the following rules:

Capacity Q proportional to speed N
Head H proportional to speed squared N^2
Horsepower P proportional to speed cubed N^3

For a series of similar fans, operating at a given point on their efficiency curves and at the same speed, performance varie according to the following rules:

Capacity Q proportional to diameter cubed D^3
Head H proportional to diameter squared D^2
Horsepower P proportional to diameter to fifth power D^5

These relations can be combined to express the performance of a family of similar fans in terms of dimensionless coefficients, as in Eqs. (1)–(3).

$$\text{Capacity coefficient } C_Q = \frac{Q}{ND^3} \tag{1}$$

$$\text{Head coefficient } C_H = \frac{gH}{N^2D^2} \tag{2}$$

$$\text{Power coefficient } C_P = \frac{gP}{wN^3D^5} \tag{3}$$

where
Q = capacity, ft^3/s
H = head, ft of fluid
P = shaft horsepower, ft-lb/s
g = gravitional acceleration, ft/s^2
N = revolutions/s
w = weight density, lb/ft^3
D = wheel diameter, ft

Values of these coefficients for a selected group of fans are listed in the **table**.

Comparison of different types of fans is further facilitated by the elimination of the diameter term. Equations for capacity coefficient and head coefficient are solved simultaneously to give specific speed N_S, as in Eq. (4). Specific speed is an inherent performance criterion; it is usually employed as a dimensionally impure coefficient. In the selection of a fan for a given application, a fan is chosen that has the desired value of N_S in its region of peak efficiency. Low specific speed corresponds to a fan of low rotative speed and large impeller diameter.

$$N_S = \frac{NQ^{1/2}}{(gH)^{3/4}} \tag{4}$$

Alternatively, speed can be eliminated, giving

Performance and dimensions of a group of fans

	Centrifugal fans			
	Backwardly curved blades	Steel plate	Forwardly curved blades	Axial fan (pressure blower)
Performance				
Specific speed (N_S)	0.25	0.11	0.21	0.75
Peak efficiency, %	75	65	65	72
Tip speed, fpm × 10^3	18	11	3.5	16
C_H at shutoff	3.8	6.3	11.2	3.2
C_Q at peak efficiency	0.42	0.18	1.2	0.4
C_P at peak efficiency	2.1	1.7	2.1	0.55
Proportions as functions of wheel diameter D				
Inlet area (×D^2)	0.7	0.4	0.8	0.8
Outlet area (×D^2)	0.6	0.3	0.7	
Axial blade length (×D)	0.3	0.4	0.6	
Number of blades	18	8	64	6

specific diameter D_S, as in Eq. (5).

$$D_S = \frac{(gH)^{0.25}D}{Q^{0.5}} \tag{5}$$

Types. The shape of a fan characteristic as a function of capacity depends on the fan type. For example, backwardly curved blade centrifugal fans have a steeply falling head characteristic and a self-limiting horsepower characteristic as shown in Fig. 3.

Fans with forwardly curved blades have rising head and horsepower characteristics. The axial fan has a falling horsepower characteristic so that at shutoff the fan may require more power than at high flow rate (Fig. 4). *See* DUCTED FAN. **Theodore Baumeister**

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; A. H. Church, *Centrifugal Pumps and Blowers*, 1944, reprint 1972; V. L. Streeter, B. Wylie, and K. W. Bedford, *Fluid Mechanics*, 9th ed., 1998.

Faraday effect

Rotation of the plane of polarization of a beam of linearly polarized light when the light passes through matter in the direction of the lines of force of an applied magnetic field. Discovered by M. Faraday in 1846, the effect is often called magnetic rotation. The magnitude α of the rotation depends on the strength of the magnetic field H, the nature of the transmitting substance, the frequency v of the light, the temperature, and other parameters. In general, $\alpha = VxH$, where x is the length of the light path in the magnetized substance and V the so-called Verdet constant. The constant V is a property of the transmitting substance, its temperature, and the frequency of the light.

The Faraday effect is particularly simple in substances having sharp absorption lines, that is, in gases and in certain crystals, particularly at low temperatures. Here the effect can be fully explained from the fundamental properties of the atoms and molecules involved. In other substances the situation may be more complex, but the same principles apply.

Rotation of the plane of polarization occurs when there is a difference between the indices of refraction n^+ for right-handed polarized light and n^- for left-handed polarized light. Most substances do not show such a difference without a magnetic field, except optically active substances such as crystalline quartz or a sugar solution. It should be noted that the index of refraction in the vicinity of an absorption line changes with the frequency (**Fig. 1***a*). *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; POLARIZED LIGHT.

When the light travels parallel to the lines of force in a magnetic field, an absorption line splits up into two components which are circularly polarized in opposite directions; that is the normal Zeeman effect. This means that, for one line, only right-handed circularly polarized light is absorbed, and for the other one, only left-handed light. The indices of refraction n^- and n^+ bear to their respective absorption frequencies the same relation as indicated in Fig. 1*a*; that is, they are identical in shape but displaced by the frequency difference between the two Zeeman components. It is evident that $n^+ - n^-$ is different from zero (Fig. 1*b*), and the magnetic rotation is proportional to this difference. The magnitude of the rotation is largest in the immediate vicinity of the absorption line and falls off rapidly as the frequency of the light increases or decreases. *See* ZEEMAN EFFECT.

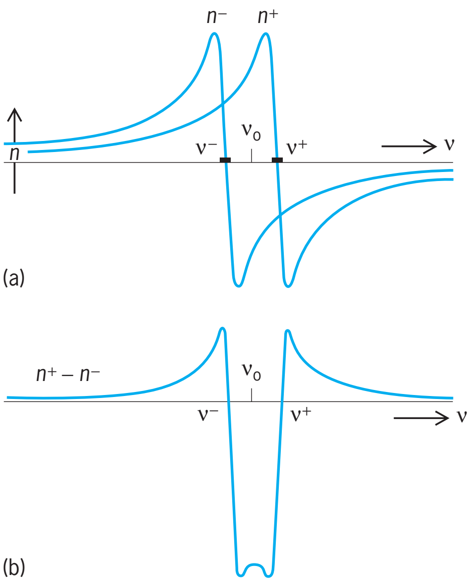


Fig. 1. Curves used in explaining the Faraday effect. (*a*) Index of refraction for left-handed circularly polarized light (n^-) and right-handed light (n^+) in the vicinity of an absorption line split into a doublet (ν^-, ν^+) in a magnetic field. (*b*) Difference between two curves, $n^+ - n^-$. Magnetic rotation is proportional to this difference.

The Faraday effect may be complicated by the fact that a particular absorption line splits into more than two components or that there are several original absorption lines in a particular region of the spectrum.

The case represented in Fig. 1 is independent of the temperature, and the rotation is symmetric on both sides of an absorption line. This case is called, not quite correctly, the diamagnetic Faraday effect. It occurs when the substance is diamagnetic, which means the splitting of the absorption line is due to the splitting of the upper level only (**Fig. 2***a*), and

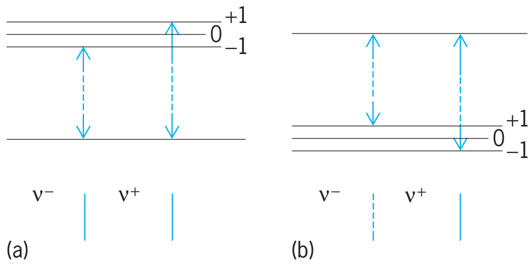


Fig. 2. Two cases of Faraday effect. (*a*) Diamagnetic case, which is temperature-independent, and (*b*) paramagnetic case, which is temperature-dependent.

the lower level of the line is not split. The same situation prevails in general when the intensity of the two Zeeman components is equal. This holds for all substances except paramagnetic salts at very low temperatures (Fig. 2*b*), in which case the v^- component is absent.

In the latter case at high temperatures there are an equal number of ions in the $+1$ and -1 states, and the two Zeeman components have equal intensities, as in the previously discussed situation. When the temperature is lowered, however, the ions concentrate more and more in the lower level (-1), and therefore absorption by ions in the upper level ($+1$) disappears. At very low temperatures only the high-frequency component in the Zeeman pattern is left. In this case the n^- refraction coefficient is not affected by the presence of the absorption line and is a constant. The difference $n^+ - n^-$ will therefore have the shape of the n^+ curve in Fig. 1*a*. Here the rotation is not symmetric with respect to the absorption line. If it is right-handed on one side of the line, it will be left-handed on the other. As the temperature is raised, the other line comes in with increasing strength until the two are nearly equal. In the transition region the Faraday effect depends strongly on the temperature. This is called the paramagnetic Faraday effect.

It is possible to modulate laser light by use of the Faraday effect with a cylinder of flint glass wrapped with an exciting coil. But since this coil must produce a high magnetic field of about 19,000 gauss, this method of modulation has been little used. For a discussion of other phenomena related to the Faraday effect *see* MAGNETOOPTICS. G. H. Dieke; W. W. Watson

Faraday's law of induction

A statement relating an induced electromotive force (emf) to the change in magnetic flux that produces it. For any flux change that takes place in a circuit, Faraday's law states that the magnitude of the emf ϵ induced in the circuit is proportional to the rate of change of flux as in expression (1).

$$\epsilon \propto -\frac{d\Phi}{dt} \qquad (1)$$

The time rate of change of flux in this expression may refer to any kind of flux change that takes place. If the change is motion of a conductor through a field, $d\Phi/dt$ refers to the rate of cutting flux. If the change is an increase or decrease in flux linking a coil, $d\Phi/dt$ refers to the rate of such change. It may refer to a motion or to a change that involves no motion.

Faraday's law of induction may be expressed in terms of the flux density over the area of a coil. The flux Φ linking the coil is given by Eq. (2), where α is

$$\Phi = \int B \cos\alpha \, dA \qquad (2)$$

the angle between the normal to the plane of the coil and the magnetic induction B. The integral is taken over the area A enclosed by the coil. Then, for a coil of N turns, Eq. (3) holds.

$$\epsilon = -N\frac{d\Phi}{dt} = -N\int \frac{d(B\cos\alpha)}{dt} A \qquad (3)$$

See ELECTROMAGNETIC INDUCTION; ELECTROMOTIVE FORCE (EMF). Kenneth V. Manning

Farm crops

The farm crops may be roughly classed as follows: (1) food crops—the bread grains (wheat and rye), rice, sugar crops (sugarbeets and sugarcane), potatoes, and dry legume seeds (peanuts, beans, and peas); (2) feed crops—corn, sorghum grain, oats, barley, and all hay and silage; and (3) industrial crops—cotton (lint and seed), soybeans, flax, and tobacco. *See* BARLEY; BEAN; CORN; COTTON; FLAX; GRAIN CROPS; OATS; PEA; PEANUT; POTATO, IRISH; POTATO, SWEET; SORGHUM; SOYBEAN; SUGAR CROPS; TOBACCO.

Regional cultivation. Crop production is regionalized in the United States in response to the combination of soil and climatic conditions and to the land topography, which favors certain kinds of crop management. In general, commercial farm crops are confined to land in humid and subhumid climates that can be managed to minimize soil and water erosion damage, where soil productivity can be kept at a relatively high level, and where lands are smooth enough to permit large-scale mechanized farm operations. In less well-watered regions, cropping is practiced efficiently on fairly level, permeable soils, where irrigation water can be supplied. The tilled crops, such as corn, sorghums, cotton, potatoes, and sugar crops, are more exacting in soil requirements than the close-seeded crops, such as wheat, oats, barley, rye, and flax. The crops planted in soil stands, mostly hay crops (as well as pastures), are efficient crops for lands that are susceptible to soil and water erosion.

United States productivity. The United States has the highest production of any area of its size in the world with regard to farm crops, together with pasture and rangelands that support livestock. Productivity per acre has increased tremendously over the years because of a combination of factors that have resulted in more efficient systems of farming. These factors include greatly improved use of mechanical and electric power and a high degree of mechanization applied to land and water management, methods of land preparation, planting, crop protection (against insects, diseases, and weeds), harvesting, curing, and storage. There have been marked advances in improving varieties of all crops and in seed technology. The development and application of effective pesticide chemicals and the greatly increased use of commercial fertilizers have been included in the newer farming systems. Soil and water conservation programs are widespread, resulting in better choice of many kinds of crops and of farming practices that are most suitable to the local climatic conditions and to the capabilities of the local classes of soils. *See* AGRICULTURAL SCIENCE (PLANT); AGRICULTURAL SOIL AND CROP PRACTICES; FERTILIZING; LAND DRAINAGE (AGRICULTURE).

Farming regions. The major farming regions of the United States are named from the predominant kinds of crops grown, even though there is tremendous diversity within each region. The Corn Belt includes a great central area extending from Nebraska and South Dakota east across much of Iowa, Missouri, Illinois, and Indiana to central Ohio. To the north and east of this region is the Hay and Dairy Region, which actually grows large quantities of feed grains. To the south of the Corn Belt is the Corn and Winter-Wheat Belt, but here also extensive acreages of other crops are grown. The southern states, once the Cotton Belt, now concentrate on hay, pasture, and livestock, with considerable acreages of soybeans and peanuts. The cultivated portions of the Great Plains, extending from Canada to Mexcio, with annual rainfall of 15–25 in. (37.5–52.5 cm) are divided into a spring-wheat region in the Dakotas and a winter-wheat region from Texas to Nebraska, with grain sorghum a major crop in all portions of the Great Plains where soil conditions and topography favor tillage. The Intermountain Region, between the Rocky Mountains and the Cascade–Sierra Nevada mountain ranges, is cropped only where irrigation is feasible, and a wide range of farm crops is grown. In the three states of the Pacific Region, a great diversity of crops is grown. Cotton is now concentrated in the irrigated regions from Texas to California.

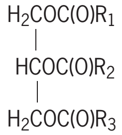
Marketing. From a world viewpoint, the United States is known most widely for its capacity to produce and export wheat. However, this nation has become a major producer of soybeans for export, and rice exports have become important. The ability to produce feed grains in abundance and at relatively low cost has created a large world market for United States corn, sorghum grains, oats, and barley. Although United States cotton once dominated the world market, its total production is now only about one-quarter of the world cotton supply. Most of the other farm crops are consumed within the United States. Tobacco and sugar crops are high acre-value crops, as are potatoes, peanuts, and dry beans. The production of these high acre-value crops is concentrated in localized areas where soils and climate are particularly favorable, rather than in broad acreages.

Howard B. Sprague

Fat and oil

Naturally occurring esters of glycerol and fatty acids that have commercial uses. Since fats and oils are triesters, they are commonly called triglycerides or simply glycerides. A glyceride may be designated a fat or oil, depending on its melting point. A fat is solid and an oil is liquid at room temperature. Some liquid waxes are incorrectly referred to as oils. *See* ESTER; TRIGLYCERIDE; WAX, ANIMAL AND VEGETABLE.

Structure and properties. The structure of triglycerides is shown below, where R_1, R_2, and R_3 represent the alkyl chain of the fatty acid.



$$\begin{array}{l} H_2COC(O)R_1 \\ \quad | \\ \ HCOC(O)R_2 \\ \quad | \\ H_2COC(O)R_3 \end{array}$$

The physical and chemical properties of fats and oils are determined to a large extent by the types of fatty acids in the glyceride. It is possible for all the acids to be identical, but this is rare. Usually there are two or even three different acids esterified to each glycerol molecule.

In all commercially important glycerides, the fatty acids are straight-chain, and nearly all contain an even number of carbon atoms. There are often traces of acids with an odd number of carbon atoms as well, but these are normally ignored. With the exception of castor oil, which contains the hydroxy substituted acid, ricinoleic acid, the alkyl chain is unsubstituted.

Most fats and oils are based on C_{16} and C_{18} acids with zero to three ethylenic bonds. There are exceptions, such as coconut oil, which is rich in shorter-chain acids, and some marine oils, which contain acids with as many as 22 or more carbons and six or more ethylenic linkages. The fatty acid composition of a number of commercial glycerides is shown in the **table**. The effect of chain length and degree of unsaturation on the chemical and physical properties is evident.

The majority of fats and oils are not important commercially. However, the noncommercial group contains the most unusual fatty acids, such as those with acetylenic bonds, epoxy rings, and cyclic structures.

Sources and isolation. The majority of fats and oils come from only a few sources. Plant sources are nuts or seeds, and nearly all terrestrial animal fats are from adipose tissue. Marine oils come principally from the whole body, although a small amount comes from trimmings.

Plant fats and oils are obtained by crushing and solvent extraction. Crushing is the older process, and while it is simpler to use than solvent extraction, it is less efficient. Sometimes high-fat-content sources are crushed or prepressed prior to solvent extraction to facilitate the extraction operation. *See* SOLVENT EXTRACTION.

Animal and marine oils are nearly all recovered by rendering. This is a process of heating fatty tissue with steam or hot water to melt and free the glyceride, followed by separating the oil or fat from the aqueous layer. The glyceride is processed for direct sale; the aqueous layer is first dried and then sold as a feed supplement. Tallow is obtained almost exclusively from beef adipose tissue; grease comes from swine. *See* ADIPOSE TISSUE.

Processing. There are some uses for fats and oils in their native state, but ordinarily they are converted to more valuable products. The most important changes are hydrolysis and hydrogenation.

Percentage fatty acid composition and some empirically determined properties of selected commercial glycerides[a]

	Saturated[b]							Unsaturated[c]										Average[d]			
Source	08	10	12	14	16	18	20	16:1	18:1	18:2	18:3	18:4	20:1	20:2	20:4	20:5	22:6	IV	SV	mp, °C	Titer, °C
Castor[e]					1	1			3	4								85	181	−11	
Coconut	8	7	48	17	8	4			5	3								8	255	24	22
Corn					11	2			27	59								124	190	−11	
Cottonseed				1	25	3			17	53								103	193	0	34
Crambe					2		2		17	9	6		61	1				91	170		
Linseed					6	4			19	15	57							175	192	−20	20
Palm				1	47	4			38									52	200	37	44
Palm kernel	1	3	51	18	9	2			15	1								19	250	25	24
Rapeseed[f]					4	1	1		17	13	5		56					81	175	−9	14
Soybean					10	3			22	55	8		1					130	193	−22	20
Tallow				3	26	22		3	43	1								50	200		42
Herring				1	14	1		10	24			11		3	9	13	6	148	190		24
Menhaden				7	24	3		10	17	1		4			3	12	9	178	191		28

[a] Values are approximate, do not include trace amounts of acids, and do not show the variation due to variety and growing conditions.
[b] 08, 10, and so on, indicate total number of carbons in the fatty acid.
[c] 18:1, 18:2, and so on, indicate total number of carbons in the fatty acid and the number of ethylenic bonds.
[d] IV = iodine value; SV = saponification value.
[e] Castor oil contains about 90% 12-hydroxy-9-octadecenoic acid.
[f] Newer varieties have less than 1% C 20:1 unsaturated acid.

Hydrolysis is commonly called splitting. The purpose is to hydrolyze the ester into its constituent glycerol and fatty acids, which are valuable intermediates for many compounds. Splitting is normally a continuous process in which the glyceride and water are fed into a tall cylindrical reactor at a temperature of about 260°C (500°F) and pressure of approximately 5200 kilopascals (750 lb/in.^{2}). Fatty acids come out the top, and sweetwater, aqueous glycerol, comes out the bottom; both are recovered and purified. *See* GLYCEROL; HYDROLYTIC PROCESSES.

Hydrogenation is the catalytic addition of hydrogen to ethylenic bonds, and is applicable to both acids and glycerides. Nickel is the customary catalyst, and the reaction is normally run at pressures to 3100 kPa (450 lb/in.^{2}) and temperatures to about 200°C (390°F). The purpose of hydrogenation is usually to raise the melting point or to increase the resistance to oxidation. *See* HYDROGENATION.

Analysis. At one time the characterization of fats and oils was exceedingly difficult due to their molecular weight and similar structure. Consequently, various empirical tests were devised to ensure quality and identity. Gas-liquid chromatography has largely supplanted these old tests, but some still remain and form the basis for commercial standards and trading rules.

Iodine value. The iodine value of a glyceride is related to its unsaturation: the higher the iodine value, the greater the unsaturation and the greater the liquidity. It is expressed in grams of iodine absorbed by 100 g of oil.

Saponification value. In this indirect procedure the ester is saponified and then back-titrated to determine the mean molecular weight of the glyceride molecule. The lower the molecular weight, the greater the saponification value. *See* DETERGENT; SOAP.

Acid value. The acid value of an oil or fat indicates the amount of nonesterified fatty acid present. The determination is made by alkaline titration.

Solidification point. This is the temperature at which a liquefied sample solidifies under certain standardized conditions. The solidification point is not necessarily the same as the melting point.

Titer. Titer is the temperature at which the first permanent cloud appears when a molten sample is cooled in a specified manner. It is an important test for detecting small amounts of contaminants.

Color. Color is almost invariably related to quality: the darker the color, the poorer the quality. There are many methods for determining color, most of which were developed for a specific product or group of products.

Gas-liquid chromatography. Gas-liquid chromatography is the most valuable analytical procedure available to the chemist of fats and oils. By selecting the proper column and operating temperatures, it is possible to obtain a quantitative estimate of the chain length, the amount of unsaturation, and the types of substitution associated with the fatty-acid chain.

Since gas-liquid chromatography requires so little sample and can be quantified, it has been used to determine, indirectly, the iodine value and the acid value of fats and oils when there has been insufficient sample for conventional analysis. Because gas-liquid chromatography is so sensitive to small amounts of trace substances, it is especially useful for distinguishing between two sources of glycerides or their derivatives. *See* CHROMATOGRAPHY; DRYING OIL; ESSENTIAL OILS; FAT AND OIL (FOOD).

Howard M. Hickman

Bibliography. R. J. Hamilton and A. Bhati (eds.), *Fats and Oils: Chemistry and Technology*, 1981; R. J. Hamilton and A. Bhati (eds.), *Recent Advances in Fats and Oils*, 1987; K. S. Markley, *Fatty Acids*:

Their Chemistry, Properties, Production, and Uses, 5 pts., 2d ed., 1983; D. Swern, *Bailey's Industrial Oil and Fat Products*, 4th ed., 2 vols., 1979–1982.

Fat and oil (food)

One of the three major classes of basic food substances, the others being protein and carbohydrate. Fats and oils are a source of energy. They also aid in making both natural and prepared foods more palatable by improving the texture and providing a more desirable flavor. *See* CARBOHYDRATE; FOOD; PROTEIN.

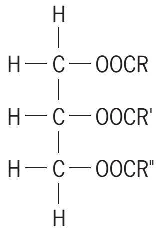
Fats and oils are esters of glycerol and fatty acids. They contain small amounts of other fat-soluble compounds—some common to all fats and others depending on the natural source. Fats and oils are soluble in organic solvents such as petroleum hydrocarbons, ether, and chloroform, but are insoluble in water. By definition, fats are more or less solid at room temperature, whereas oils are liquid. Some tropical oils, such as coconut oil, are solid at normal room temperature, but are liquid in their natural habitat.

Fats are grouped according to source. Animal fats are rendered from the fatty tissues of hogs, cattle, sheep, and poultry. Butter is obtained from milk. Vegetable oils are pressed or extracted from various plant seeds, primarily from soybean, cottonseed, corn (germ), peanut, sunflower, safflower, olive, rapeseed, sesame, coconut, oil palm (pulp and kernel separately), and cocoa beans. Marine oils are not consumed in the United States, but commonly are elsewhere. They are obtained mostly from herring, sardine, and pilchard.

Nutritive value. Fats and oils are important in the diet. They are the most concentrated form of food energy, contributing about 9 cal/g (38 joules/g), as compared to about 4 cal/g (17 joules/g) for carbohydrates and proteins. Fats make a meal more satisfying by creating a feeling of fullness, and also delay the onset of hunger. Contrary to popular belief, fats are highly digestible, with 94–98% of the ingested fat being absorbed from the intestinal tract.

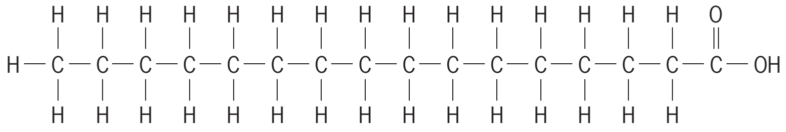
The polyunsaturated fatty acids, primarily linoleic and arachidonic, are essential nutrients; that is, they are not synthesized by the body but are required for tissue development. Absence of these fatty acids from the diet results in an essential fatty acid syndrome and in a specific form of eczema in infants. Vegetable oils are an excellent source of linoleic acid, while meat fats provide arachidonic acid in small but significant amounts. Fats and oils are carriers of the oil-soluble vitamins A and D, and are the main source of vitamin E. They also have a sparing action on some of the B complex vitamins. *See* NUTRITION; VITAMIN.

Chemical constitution. Fats and oils, also called triglycerides, are esters of the trihydric alcohol glycerol, $C_3H_5(OH)_3$, and various fatty acids (see **table**). Most naturally occurring fatty acids are straight carbon chains with a terminal carboxyl group and 4–24 carbon atoms in even numbers. A few acids have an odd number of carbons and some have a cyclic group or branched chain, but all of these are relatively rare. Triglyceride structure is usually represented as shown in structure (**1**), where R, R′, and R″ represent carbon chains of the fatty acids. *See* TRIGLYCERIDE.

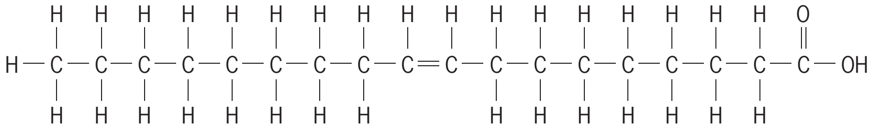

$CH_2(OOCR)-CH(OOCR')-CH_2(OOCR'')$

(**1**)

The fatty acid may be saturated or unsaturated. In a saturated acid, such as palmitic acid (**2**), all bonds between carbon atoms are single bonds, with hydrogen atoms attached to all the carbon atoms, except that of the carboxyl radical. In an unsaturated acid, each of two adjacent carbons lacks one hydrogen, so that the carbon bonds link together to form a double bond. Oleic acid (**3**), with one double bond, is a monounsaturated acid. Linoleic, linolenic, and arachidonic acids are polyunsaturated, having two, three, and four double bonds, respectively.


$CH_3(CH_2)_{14}COOH$

(**2**)


$CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$

(**3**)

Typical fatty acid composition of common fats and oils[a]

	Saturated			Unsaturated		
Type	Palmitic	Stearic	Other[b]	Oleic	Linoleic	Other
Lard	24.8	12.3	2.6	45.1	9.9	5.3[c]
Beef fat	25.5	21.6	6.5	38.7	2.2	5.5[c]
Butterfat	26.2	12.5	26.5	28.2	2.9	3.7[c]
Palm oil	45.1	4.7	1.6	38.8	9.4	0.4
Cocoa butter	25.8	34.5	1.2	35.3	2.9	0.3
Soybean	11.0	4.0	0.5	23.4	53.2	7.9[d]
Cottonseed	24.7	2.3	1.1	17.6	53.3	1.0[e]
Corn	12.2	2.2	0.1	27.5	57.0	1.0
Peanut	11.6	3.1	5.7	46.5	31.4	1.7
Sunflower	6.8	4.7	1.1	18.6	68.2	0.6
Olive	13.7	2.5	0.9	71.1	10.0	1.8
Coconut	8.4	2.5	81.1	6.5	1.5	0.0

[a] As weight percentages of component fatty acids.
[b] Butterfat and coconut oil contain saturated fatty acids having 4–14 and 6–14 carbon atoms, respectively. Peanut oil contains fatty acids with 20–24 carbon atoms.
[c] Mainly palmitoleic acid and about 0.4% arachidonic acid.
[d] Mostly linolenic acid.
[e] Includes 0.5–1% malvalic and sterculic acids with a terminal cyclopropene ring structure.

Unsaturated acids can be converted to saturated acids by the addition of hydrogen.

Chain length is also important. The most common saturated fatty acids in edible fats and oils are palmitic (16 carbons), stearic (18), and lauric (12). The important unsaturated acids have 18 carbons, except for arachidonic with 20.

The physical characteristics of fat depend on the distribution of fatty acids on the various triglyceride molecules. The distribution is complex and nonrandom. Lard, for example, has palmitic acid almost exclusively on the second or middle carbon atom of glycerine. Vegetable oils have unsaturated fatty acids in this position, with saturates exclusively on the end carbons of glycerine. Cocoa butter consist almost exclusively of 2-oleo-palmitostearin and 2-oleodistearin.

Most fats contain 0.5–2% nonglyceride components which are classified as unsaponifiables, including squalene, carotenoids, tocopherols, and sterols. Tocopherols, also known as vitamin E, have antioxidant activity. Cholesterol is a sterol found only in animal fats. Plant sterols are chemically related to cholesterol, but are physiologically inactive in humans, except that they may competively inhibit absorption of cholesterol. *See* CHOLESTEROL; STEROL; VITAMIN E.

Production methods. Processing of fats and oils is carried out in a series of individual steps: extraction, refining, bleaching, and deodorization. Additional processes for specific products include winterization, hydrogenation, and texturizing. Typically there are many variations in both equipment and technique for each process. The selection of equipment also depends on the oilseed or fatty tissue to be processed.

Fats and oils are contained in seed or animal tissues within proteinaceous cell walls. The proteins are coagulated, usually through the use of heat, to release the oils. Animal tissues need no prior preparation other than cutting or comminuting them. Oilseeds are prepared in a variety of ways before oil can be removed: cottonseed must be dehulled; soybeans may or may not be dehulled; sunflower seeds are rarely dehulled; coconuts, palm kernels, and peanuts must have their shells removed; corn germ must be separated from the kernel before its oil can be removed. *See* CORN; COTTON; PEANUT; SOYBEAN; SUNFLOWER.

Fat rendering. Beef fat or tallow is usually dry-rendered under vacuum. The cracklings are then strained from the bulk of the fat and pressed to remove more tallow. Prime steam lard is wet-rendered by heating hog fat under 40–60 lb/in.^{2} (276–414 kilopascals) pressure by using steam injection into the vessel. After cooking for 4–6 h, the mixture is allowed to settle. The separated lard is then drawn off. In low-temperature rendering, comminuted fatty tissues are heated to 115–120°F (46–49°C) to melt the fat, which is separated by centrifugation. The tissue residues may be used in edible meat products, such as sausage.

Oil extraction. Oil is removed from oilseeds by pressing, solvent extraction, or a combination of both. First, however, the seeds must be crushed or flaked and cooked to denature the proteins. The moisture content must also be adjusted to an optimum level for the particular seed and extraction method. Some mills use high pressure presses which also develop high temperatures in the oil and cause some degradation. Extraction with hexane, the most commonly used solvent, gives the best-quality oil, but is not practical for all oilseeds. Prepressing seeds at low pressure (and consequently low temperature) and removing the balance of the oil with solvent is frequently the best compromise.

Oil refining. Crude oils received from the mill must be refined for use. The first step in the process is called refining; the term "fully refined" refers to an oil which has been refined, bleached, and deodorized. Refining as such is carried out by mixing the crude oil with a water solution of sodium hydroxide at about 150°F (66°C). The alkali reacts with the free fatty acids to form soap, which is insoluble in oil. Mucilaginous gums, oil soluble in dry form, are hydrated and become insoluble. Other minor components, such as part of the sterols and tocopherols, are also incorporated in the insoluble "soap stock." The entire mass is pumped through a continuous centrifuge to remove the oil-insoluble matter. The oil is then water-washed, recentrifuged, and dried under vacuum. *See* CENTRIFUGATION.

Some vegetable oils, especially soybean oil, may first be degummed by using water to hydrate the gums. The crude gums which separate from the oil are dried to yield crude lecithin, which is useful as an emulsifier. The degummed oil is then alkali-refined as before.

Animal fats and water- or phosphoric acid–degummed vegetable oils may be steam-refined, stripping free fatty acids, monoglycerides, unsaponifiable waxes, and some pigments. During this process, the oil is held under high vacuum (4–6 torr or 0.5–0.8 kPa) and high temperature (about 450°F or 230°C).

Bleaching. Bleaching, the process for removing pigments from fats and oils, is usually carried out by adding about 1% bleaching clay to oil under vacuum at about 225°F (107°C), agitating, and then filtering to remove the clay. High temperature drives moisture from the clay so that it will adsorb the pigments. Neutral clay removes most pigments, but acid-activated earth also removes chlorophyll. Carotene, found in palm oil and tallow, is not removed by adsorption but can be decolorized by heat bleaching during deodorization. Some pigments are fixed in the oil, and cannot be removed.

Deodorization. Deodorization is used to remove volatile materials from the oil product. Along with removal of compounds which contribute flavor and odor, free fatty acids, monoglycerides, and some color bodies are distilled off. Peroxides and carotenes are decomposed. The resulting product is bland in flavor and odor. The process is carried out by blowing steam through the oil held at 400–500°F (200–260°C) and a vacuum of 5–6 torr (0.7–0.8 kPa).

Fractional crystallization. Fractional crystallization of fats is the separation of the solidified fat into hard fractions (stearine) and soft fractions (oil). It is carried out by allowing the fat to crystallize at a preselected temperature and filtering or pressing the mass to separate the oil from the stearine. The process is used for preparing high-stability liquid frying oils from partially hydrogenated soybean oil and from palm oil.

Hard butters are made by fractionating hydrogenated oils or naturally hard fats from solvent to isolate the disaturated, monounsaturated triglycerides, which resemble cocoa butter in melting characteristics.

Winterization. Winterization is a specialized form of fractional crystallization of fats. Natural cottonseed oil and partially hydrogenated soybean oil contain solid fats at low to moderate room temperature. This makes these oils undesirable in appearance, especially when refrigerated, and useless for mayonnaise, where the fat crystals would break the emulsion. Winterization is achieved by holding the oil at 40°F (4°C) for 2–4 days to crystallize the solid fats, which are removed by filtration. Crystal inhibitors, primarily oxystearin and polyglycerol esters of fatty acids, may be added to the clear oil to retard further crystallization of fats, which are rarely completely removed by winterization. Some oils, particularly from sunflower and corn, have a high level of wax, which causes a cloudy appearance. These oils are dewaxed in a process resembling winterization.

Hydrogenation. Hydrogenation is the chemical binding of hydrogen to the double bonds of unsaturated fatty acids. Glyceride molecules with two or more unsaturated acids are liquids at room temperature. Saturating one or more of these acids makes the glyceride more solid and also more stable against oxidation. The hydrogenation process calls for vigorous agitation of a mixture of oil, finely divided nickel metal catalyst (0.025–0.3%), and hydrogen gas at high temperature (usually 250–400°F or 120–200°C) and gas pressure ranging from atmospheric (14.7 lb/in.^{2} or 101 kPa) to 60 lb/in.^{2} (414 kPa). The end point is measured by the refractive index, which correlates with the iodine value, and is predetermined by experience to give the desired hardness of product. Characteristics of the hardened oil depend on process temperature, type and amount of catalyst, purity and pressure of hydrogen, and other similar variables. The fat is rebleached and filtered to remove the catalyst as a final step. *See* HYDROGENATION.

Interesterification. Interesterification is used to rearrange a nonrandom mixture of fatty acids on various triglyceride molecules into a randomized distribution. It is used mainly to change lard from a coarse crystalline fat to a smooth, textured shortening which will cream in bakery products, and to convert a mixture of fully hydrogenated coconut or palm oils and fully hydrogenated cottonseed or other nonlauric oils to hard butters for confectionery use. Interesterification involves heating the fatty starting material with a catalyst, usually sodium metal or sodium methoxide, at 95–135°F (35–57°C) for

0.5–1 h, killing the catalyst with water, and removing the resulting soaps.

Texturizing. Texturizing of shortening compounds, usually a blend of partially hydrogenated oils or interesterified lard and 8–12% fully hydrogenated cottonseed or palm oils, is carried out by rapid cooling in a scraped-surface internal-chilling machine and whipping the cold, partially crystallized fat with air or nitrogen. The shortening is filled into containers to harden on standing. The process is analogous to the freezing of ice cream.

Chemical adjuncts. A number of additives are used to modify the performance characteristics of edible oils and shortenings. Emulsifiers are the largest and most varied group of additives. Monoglycerides are the basic emulsifiers for incorporation of air into cakes and icings and for retarding staling in yeast-raised doughs. These attributes are enhanced by the addition of lactylated or ethoxylated monoglycerides, polyglycerol esters of fatty acids, propylene glycol monostearate, and polysorbates. Lecithin, a very potent emulsifier, is useful as an antisticking and antispattering agent in grill frying oils and margarine and as a texture modifier in chocolate and confectionery coatings.

Acetylated monoglycerides have some emulsifying properties, but are more useful as stable oils or waxes for coating meat products, cheese, nuts, raisins, cereals, and so forth to retard moisture transfer during storage.

Antioxidants are used to retard oxidative rancidity in fats and oils. Tertiary butyl hydroquinone (TBHQ), butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT), and propyl gallate are the most widely used antioxidants. Citric acid is used as a trace-metal scavenger to inactivate the prooxidant effect of heavy metals, primarily iron and copper. Citric acid is added without exception to all oils in the final stages of deodorization. *See* CITRIC ACID.

Methyl silicones are added to some fats used in deep fat frying as an antifoam agent.

Margarine, popcorn oil, pan frying oils, and some shortenings may be colored yellow through the addition of carotene or annatto pigments. Buttery flavors are also added to margarine and popcorn oil to enhance their appeal. *See* MARGARINE.

Testing procedures. Crude oils are purchased according to rigid specifications which include free fatty acid, refining loss, bleached color, and peroxide value. Free fatty acid is found by titrating the oil with standard alkali. Refining loss is done either by cup refining, a laboratory simulation of plant refining, or by adsorption of impurities from the oil by using powdered aluminum oxide. Refined oil is bleached with clay in the laboratory for determination of color by visual comparison with a series of color standards. The most common standards are Lovibond glasses in graduated steps of yellow and red. Peroxide value, a measure of the extent of oxidative deterioration of the oil, is determined by iodine titration, in which the peroxide equivalent is found in terms of elemental iodine liberated from potassium iodide. Other evaluations of crude oil include moisture, insoluble impurities (meal or crackling residues), and volatiles such as residual solvents.

Finished salad oils must pass a cold test: the length of time the oil takes to show the first trace of fat crystals in an ice bath. The standard test requires the oil to be clear after 5.5 h. Iodine value, the amount of iodine reacting with the oil, determines the total unsaturation. Iodine values are calculated from the fatty acid composition of the oil as determined by gas-liquid chromatography. *See* GAS CHROMATOGRAPHY.

Naturally hard fats and hydrogenated vegetable oils are actually mixtures of solid and liquid triglycerides. The solid fat index (SFI) is a measure of the solid glyceride levels at various temperatures below the melting point of the fat. It is a valuable tool for determining the potential firmness of a plasticized shortening or margarine. The consistency of these products is evaluated after plasticizing by observing package penetration. A standardized needle or other pointed object is allowed to penetrate the product for several seconds. The depth of penetration must fall within predetermined limits for the material involved. Hardness is also given in terms of melting point. Fats do not have a sharp melting point, so that several empirical methods have been developed, each giving slightly different values from the others.

Potential stability of fats toward oxidation is usually determined by an accelerated test known as the active oxygen method (AOM), in which stability is measured by bubbling air through the sample at a specified temperature, usually 208°F (98°C), and determining the length of time the fat takes to reach a peroxide value of 100. Smoke point, the temperature at which a fat gives off a steady stream of smoke, is determined for frying fats. Monoglyceride content is an important evaluation for bakery shortenings. Finished oils are also evaluated for peroxide value, free fatty acid content, moisture, color, and flavor.

Many other tests exist to give special information where needed, but are not in general use.

Deterioration factors. Several forms of deterioration may occur in fats and oils.

Flavor reversion. Flavor and odor may develop after deodorization of a product to complete blandness. The flavor is generally characteristic of the oil source and is therefore usually acceptable. However, soybean oil can develop disagreeable flavors described as beany, grassy, painty, fishy, or like watermelon rind. Beef fat can become tallowy, which is also objectionable. Reversion is apparently caused by changes in substances which have been oxidized prior to, but not removed by, deodorization. Ordinary chemical tests do not indicate any change in the oil. It is strictly an organoleptic observation, although work using gas chromatography has demonstrated that a large number of volatile compounds form during development of reverted flavor. Those compounds which have been identified indicate that reversion is related to oxidation of unsaturated fatty acids.

Oxidative rancidity. This is a serious flavor defect and highly objectionable. It starts with the formation of hydroperoxides at the double bond of fatty acids, primarily linolenic and linoleic, which then decompose to form aldehydes which have a pungent, disagreeable flavor and odor. Oxidative rancidity is detectable chemically by peroxide value and by one of several tests for aldehydes. Peroxides and aldehydes in crude oils are removed by deodorization. They are reformed in finished oils by further oxidation, which is catalyzed by exposure of the oil to light, to some metals, especially copper and iron, and to excessive heat. Retardation of oxidation is brought about by using opaque, airtight containers, or nitrogen blanketing if clear glass bottles are used. Antioxidants, which retard oxidation by interrupting the reaction of oxygen radicals, are required in meat fats, since lard, tallow, and so on contain no natural antioxidant material. Vegetable oils contain tocopherols. Additional antioxidant, with the exception of TBHQ, has little benefit for these oils. Copper metal and copper-bearing alloys are extremely active prooxidants, and must never be used in oil-handling or oil-processing equipment. Iron rust is a prooxidant, but recently cleaned iron, coated with oil to prevent rusting, is inert, and is commonly used for oil-processing and oil-storage equipment.

Hydrolytic rancidity. This type of rancidity results from the liberation of free fatty acids by the reaction of fats and oils with water. While most fats show no detectable off flavors, coconut and other lauric acid oils develop a soapy flavor, and butter develops the strong characteristic odor of butyric acid. In some foods, active lipase from materials such as raw nuts or low-temperature–pasteurized milk or from contaminating microorganisms catalyzes the hydrolysis of fats and oils. Packaged coconut-oil products and lauric-type hard butters sometimes contain added lecithin, which acts as a moisture scavenger, thereby retarding hydrolytic rancidity development.

Frying fat breakdown. Fats and oils used in deep fat frying can break down under adverse conditions, especially where frying is intermittent or the fryer capacity is not fully used. This results in a low fat turnover rate, that is, an insufficient percentage of fresh fat being added to the fry kettle at regular intervals. The breakdown is detected first by an increase in acidity, measured by free fatty acid determination but actually caused by development of acidic breakdown products. Further deterioration results in the oil becoming very dark in color, viscous, foul-odored, and foaming badly during frying. It becomes oxidized and then polymerized, requiring that it be discarded, since it imparts strong off flavors to the fried food.

Crystal transformation. Crystal structure transformation of packaged shortening results in formation of a grainy, soft product, which may also lose incorporated gas and take on the appearance of petroleum jelly. In extreme cases, liquid oil pockets form in the shortening mass. In addition, there is a loss of creaming ability in baking of cakes and icing preparation. Crystal transformation is caused by improper hydrogenation, poor formulation, unstable hard fat used for texturizing, or the wrong chilling conditions. Unfortunately, such transformation takes several days after filling the shortening before it becomes obvious. Similar changes in crystal structure can cause bloom in chocolate coatings, a defect which gives a white haze or even open grain on an originally smooth, glossy surface. Chocolate bloom can be inhibited by the addition of lecithin or polysorbates or both. *See* COCOA POWDER AND CHOCOLATE; FOOD MANUFACTURING; LIPID.

Theodore J. Weiss

Bibliography. H. J. Deuel, Jr., *The Lipids: Their Chemistry and Biochemistry*, 3 vols., 1951–1957; E. W. Eckey, *Vegetable Fats and Oils*, 1954; A. Kramer and B. A. Twigg (eds.), *Quality Control for the Food Industry*, vol. 2, 3d ed., 1972; V. C. Mehlenbacher, *The Analysis of Fats and Oils*, 1960; *Official and Tentative Methods of the American Oil Chemists' Society*, 1946–1983; D. Swern, *Bailey's Industrial Oil and Fat Products*, 4th ed., vol. 1, 1979, vol. 2, 1982; T. J. Weiss, *Food Oils and Their Uses*, 2d ed., 1983.

Fate maps (embryology)

Diagrams of embryos showing what will happen to each small region in the course of development. These diagrams show where the cells move and what structures they develop into. *See* EMBRYOGENESIS.

Animals. In animals, fate maps are usually constructed by applying small patches of label to the embryo and locating the position of the labeled cells at subsequent times during development (**Fig. 1**). In the early twentieth century, vital dyes or fine carbon particles were used for labeling. Vital dyes are stains that can color living tissues without harming them, although they do tend to diffuse and fade. In the late twentieth century, the carbocyanine dyes diI and diO were introduced; they are intensely fluorescent and are retained within the membranes of cells.

Fate maps may also be constructed by grafting small explants of tissue from a labeled embryo to the same position in an unlabeled host and observing the movements and later development of the cells of the graft. The donor for such a graft would usually be an embryo carrying a marker gene such as the gene for green fluorescent protein. There are also lineage labels, such as the enzyme horseradish peroxidase or fluorescent derivatives of the polysaccharide dextran, that can be injected into cells and do not diffuse out again. These may be used either for labeling the donor embryos for grafts or to label individual cells by injection.

A combination of these methods has enabled accurate fate maps to be produced for the types of embryos most often used in biological research: those of the amphibian (**Fig. 2**), chick, mouse, fruit fly,

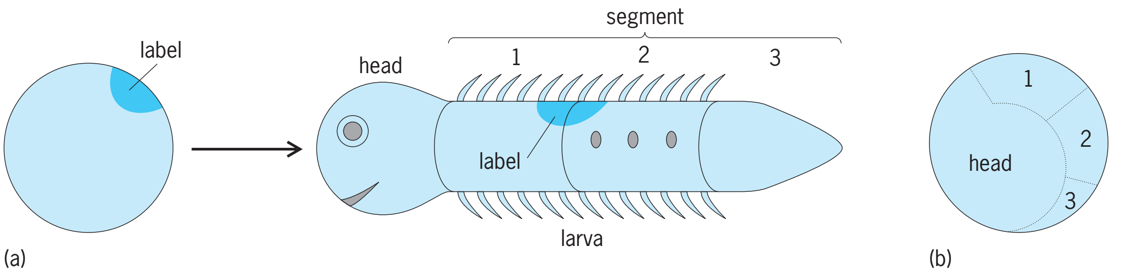


Fig. 1. Principle of fate mapping. (*a*) Label placed in a certain position on the egg ends up in a reproducible position on the later animal. (*b*) Possible fate map for this type of egg. (*Reprinted with permission from J. M. W. Slack, Essential Developmental Biology, 2d ed., Blackwell Science, 2005*)

and zebrafish. For embryos in which there is no increase in size during development and no random mixing of cells, the fate map can be precise from the fertilized egg onward. This is the case for some invertebrates such as gastropods or ascidians. For embryos in which there is some local cell mixing, as in vertebrates, the fate map cannot be so precise, and represents the average behavior of a population of embryos.

Although it is a common misconception, a fate map of multicellular embryo regions does not, in itself, give any information about the intrinsic developmental commitment of the cells of the embryo to form particular structures or cell types. For example, in cases in which the commitment of a region is acquired in response to an inductive signal from another region, the commitment may differ from the fate shown on the fate map at the time before the inductive signal operates. States of commitment must be deduced by other types of experiment such as culturing embryo regions in isolation or grafting them to different positions in a host embryo.

An associated concept to the fate map is the cell lineage, which shows the family tree of cells descended from a single precursor cell. Unlike a fate map, a cell lineage does not show the spatial relationships between cells. The lineage of every cell has been established from the fertilized egg to the adult in the nematode *Caenorhabditis elegans*, in which every individual embryo undergoes exactly the same sequence of cell divisions.

The term clonal analysis is used to refer to the deduction of developmental mechanisms from the study of cell lineage. For example, if one labeled cell predictably forms a whole structure in the later organism, then it must be the sole precursor to that structure. Conversely, if one labeled cell gives rise to progeny that span two structures, then it could not have been committed to form either structure at the time of labeling. *See* CELL LINEAGE; DEVELOPMENTAL BIOLOGY; EMBRYOLOGY. J. M. W. Slack

Plants. In clonal analysis of plants, cells are labeled by exposing the seed or developing plant to ionizing irradiation or a chemical mutagen to produce chromosome mutations that result in distinct phenotypic alterations, usually deficiencies in the pigments chlorophyll or anthocyanin. The low frequency of mutations produced indicates that these are single-cell events. Because all of the progeny of a labeled cell will carry the same chromosome mutation, a shoot meristem cell labeled with a chlorophyll deficiency mutation, for example, will produce a sector, or clone, of white tissue in the developing plant (**Fig. 3**). It is possible to deduce the number and fate of meristem cells labeled at one stage of development by examining the size and position of sectors present in the plant at a later stage of development. For example, a sector extending the entire length of the shoot would be generated by a permanent initial cell at the center of the meristem; if the width of

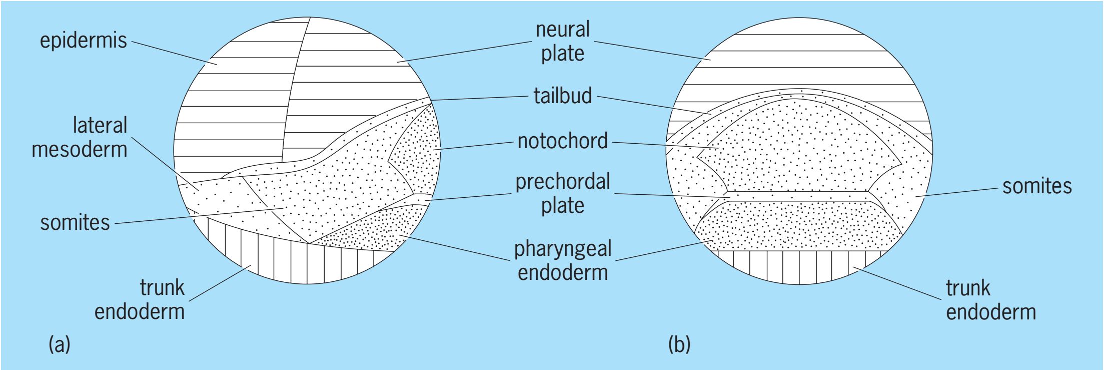


Fig. 2. Fate maps of urodele blastula from (*a*) lateral and (*b*) dorsal aspects, deduced from vital staining. (*After J. M. W. Slack, From Egg to Embryo: Determinative Events in Early Development, Cambridge University Press, 1983*)

that sector occupied one-third of the circumference of the shoot, then the shoot must have been formed by three such initial cells. *See* MUTATION; PLANT GROWTH.

Fate maps have been constructed for the shoot apical meristems of maize (*Zea mays*) and sunflower (*Helianthus annuus*) embryos based on pigment-deficient sectors generated in plants grown from irradiated seed (**Fig. 4**). The pattern of sectors in both plants indicates that several domains of cells from the periphery to the center of the meristem give rise to successively more apical portions of the mature plant. It also shows that the contribution of cells occupying similar positions in the meristem is variable. Thus the fate of cells in the meristem of these two plants is not fixed in the embryo: no cells are committed to a particular developmental fate, such as the production of the reproductive parts of the plant. The number of cells and cell domains predicted by the fate maps of maize and sunflower is consistent with the actual number and arrangement of cells observed in histological sections of the meristems.

Similarly, the fate of cells in developing leaf primordia of tobacco (*Nicotiana tabacum*) has been established by irradiating primordia at different developmental ages. A different type of pigmentation chimera that was produced by certain interspecific crosses has been used to construct a fate map of cells contributing to the cotyledons, to the first true leaves, and to the shoot meristem in the developing cotton (*Gossypium*) embryo. Earlier, the fate of cells in the different cell layers of the shoot meristem of several plants was deduced from studies of chimeras in which different cell layers of the meristem are genetically distinct for pigmentation or chromosome number. That research showed that whereas cells in a particular layer of the meristem generally contribute to a particular part of the shoot, such as the outermost layer to the epidermis or the third layer to the vascular tissue and pith, the cell layers do not have fixed developmental fates. *See* CHIMERA; DEVELOPMENTAL BIOLOGY; EMBRYOLOGY. Dorothy E. Jegla

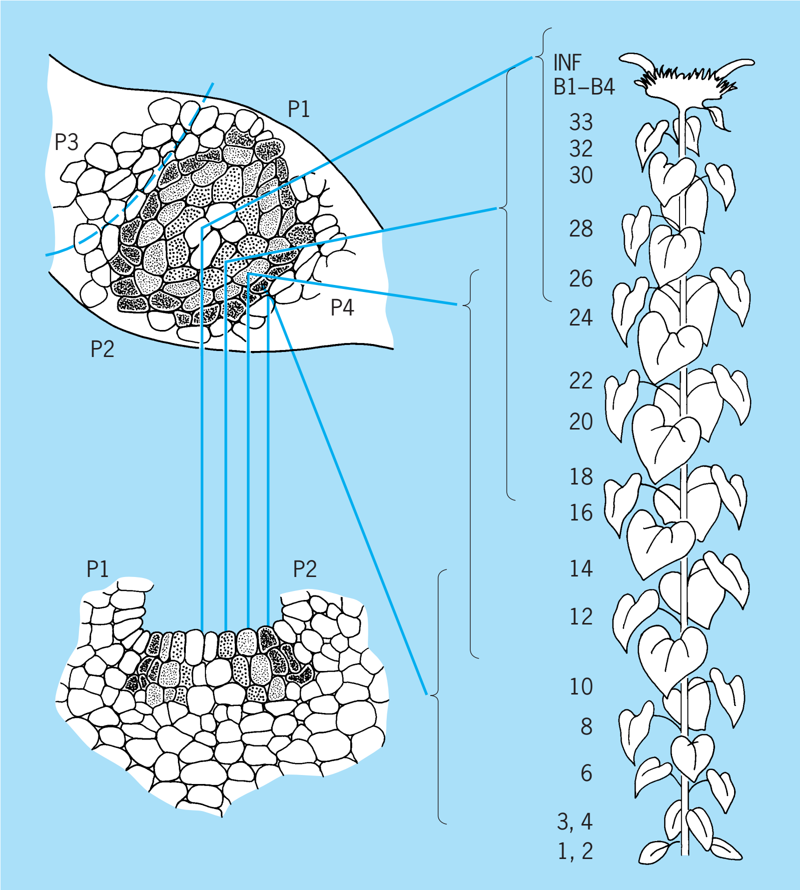


Fig. 4. Fate map of the shoot apical meristem of the sunflower embryo. Cell drawings on the left show the actual number and arrangement of cells in the surface layer (top) and median longitudinal section (bottom) of a typical meristem. Positions of the first four leaves that are present in the embryo are labeled P1–P4. Cells are differentially shaded to show four domains of cells that give rise to the successively more apical portions (indicated by brackets) of the mature plant diagrammed on the right. Numbers refer to node number: B1–B4 to the involucral bracts, and INF to the inflorescence. (*After D. E. Jegla and I. M. Sussex, Cell lineage patterns in the shoot meristem of the sunflower embryo in the dry seed, Dev. Biol., 131:215–225, 1989*)

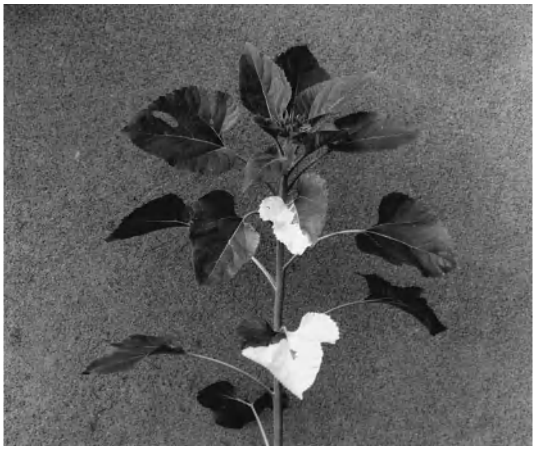


Fig. 3. Chlorophyll-deficient sector in a sunflower plant grown from irradiated seed. The sector represents a clone of tissue derived from a single cell present in the meristem of the embryo at the time of irradiation.

Bibliography. D. E. Jegla and I. M. Sussex, Cell lineage patterns in the shoot meristem of the sunflower embryo in the dry seed, *Dev. Biol.*, 131:215–225, 1989; J. M. W. Slack, *From Egg to Embryo: Regional Specification in Early Development*, 2d ed., Cambridge University Press, 1991; J. M. W. Slack, *Essential Developmental Biology*, 2d ed., Blackwell Science, 2005.

Fault analysis

The detection and diagnosis of malfunctions in technical systems. Such systems include production equipment (chemical plants, steel mills, paper mills, and power stations), transportation vehicles (ships, airplanes, automobiles), and household appliances (washing machines, air conditioners). In any of these systems, malfunctions of components may lead to damage of the equipment itself, degradation of its function or product, jeopardy of its mission, and hazard to human life. While the need to detect and diagnose malfunctions is not new, advanced fault detection has been made possible only by the proliferation of the computer. Fault detection and diagnosis actually means a scheme in which a computer monitors the technical equipment to signal any malfunction and determines the components responsible. The detection and diagnosis of the fault may be followed by automatic actions, enabling the fault to be corrected such that the system may operate successfully even under the particular faulty condition.

Diagnostic concepts. Fault detection and diagnosis applies to both the basic technical equipment and the actuators and sensors attached to it. In the case of a chemical plant, the former includes the reactors, distillation columns, heat exchangers, compressors, storage tanks, and piping. Typical faults are leaks, plugs, surface fouling, and broken moving parts. The actuators are mostly valves, together with their driving devices (electric motors and hydraulic or pneumatic drives). The sensors are devices measuring the different physical variables in the plant, such as thermocouples, pressure diaphragms, and flow meters. Actuator and sensor fault detection is very important because these devices are prone to faults.

The on-line or real-time detection and diagnosis of faults means that the equipment is constantly monitored during its regular operation by a permanently connected computer, and any discrepancy is signaled almost immediately. On-line monitoring is very important for the early detection of any component malfunction before it can lead to more substantial equipment failure. In contrast, off-line diagnosis involves monitoring the system by a special, temporarily attached device, under special conditions (for example, car diagnostics at a service station).

The diagnostic activity may be broken down into several logical stages. Fault detection is the indication of something going wrong in the system. Fault isolation is the determination of the fault location (the component which malfunctions), while fault identification is the estimation of its size. On-line systems usually contain the detection and isolation stage; in off-line systems, detection may be superfluous. Fault identification is usually less important than the two other stages.

Fault detection and isolation can never be performed with absolute certainty because of circumstances such as noise, disturbances, and model errors. There is always a trade-off between false alarms and missed detections, with the proper balance depending on the particular application. In professionally supervised large plants, false alarms are better tolerated and missed detections may be more critical, while in consumer equipment (including cars) the situation may be the opposite.

Approaches. A number of different approaches to fault detection and diagnosis may be used individually or in combination.

Limit checking. In this approach, which is the most widely used, system variables are monitored and compared to preset limits. This technique is simple and appealing, but it has several drawbacks. The monitored variables are system outputs that depend on the inputs. To make allowance for the variations of the inputs, the limits often need to be chosen conservatively. Furthermore, a single component fault may cause many variables to exceed their limits, so it may be extremely difficult to determine the source. Monitoring the trends of system variables may be more informative, but it also suffers from the same drawbacks as limit checking.

Special and multiple sensors. Special sensors may be applied to perform the limit-checking function (such as temperature or pressure limit sensors) or to monitor some fault-sensitive variable (such as vibration or sound). Such sensors are used mostly in noncomputerized systems. Multiple sensors may be applied to measure the same system variable, providing physical redundancy. If two sensors disagree, at least one of them is faulty. A third sensor is needed to isolate the faulty component (and select the accepted measurement value) by "majority vote." Multiple sensors may be expensive, and they provide no information about actuator and plant faults.

Frequency analysis. This procedure, in which the Fourier transforms of system variables are determined, may supply useful information about fault conditions. The healthy plant usually has a characteristic spectrum, which will change when faults are present. Particular faults may have their own typical signature (peaks at specific frequencies) in the spectrum. *See* FOURIER SERIES AND TRANSFORMS.

Fault-tree analysis. Fault trees are the graphic representations of the cause–effect relations in the system. On the top of the tree, there is an undesirable or catastrophic system event (top event), with the possible causes underneath (intermediate events), down to component failures or other elementary events (basic events) that are the possible root causes of the top event. The logic relationships from bottom up are represented by AND and OR (or more complex) logic gates. Fault trees can be used in system design to evaluate the potential risks associated with various component failures under different design variants (bottom-up analysis). In a fault diagnosis framework, the tree is used top down; once the top event is observed, the potential causes are analyzed by following the logic paths backward.

Parameter estimation. This procedure uses a mathematical model of the monitored system. The parameters of the model are estimated from input and output measurements in a fault-free reference situation. Repeated new estimates are then obtained on-line in the normal course of system operation. Deviations from the reference parameters signify changes in the plant and a potential fault. The faulty component location may be isolated by computing the new physical plant parameters and comparing them with those from the model. *See* ESTIMATION THEORY; MODEL THEORY.

Consistency checking. This is another way of using the mathematical-system model. The idea is to check if the observed plant outputs are consistent with the outputs predicted by the model (**Fig. 1**). Discrepancies indicate a deviation between the model and the plant (parametric faults) or the presence of unobserved variables (additive faults). This testing concept is also called analytical redundancy since the model equations are used in a similar way as multiple sensors.

In preparation for fault monitoring by analytical redundancy methods, a mathematical model of the plant needs to be established. This may be done from "first principles," relying on the theoretical understanding of the plant's operation, or by systems identification using experimental data from a fault-free plant.

The actual implementation of fault monitoring usually consists of two stages (Fig. 1). The first is residual generation, where residuals are mathematical quantities expressing the discrepancy between the actual plant behavior and the one expected based on the model. Residuals are nominally zero and become nonzero by the occurrence of faults. The second stage is residual evaluation and decision making, where the residuals are subjected to threshold tests and logic analysis. Disturbances and model errors may also cause the residuals to become nonzero, leading to false alarms.

Fault isolation requires specially manipulated sets of residuals. In the most frequently used approach, residuals are arranged so that each one is sensitive to a specific subset of faults (structured residuals). Then in response to a particular fault, only a fault-specific subset of residuals triggers its test, leading to binary fault codes.

Principal component analysis (PCA). In this approach, empirical data (input and output measurements) are collected from the plant. The eigenstructure analysis of the data covariance matrix yields a statistical model of the system in which the eigenvectors point at the "principal directions" of the relationships in the data, while the eigenvalues indicate the data variance in the principal directions. This method is successfully used in the monitoring of large systems. By revealing linear relations among the variables, the dimensionality of the model is significantly reduced. Faults may be detected by relating plant observations to the normal spread of the data, and outliers indicate abnormal system situations. Residuals may also be generated from the principal component model, allowing the use of analytical redundancy methods in this framework. *See* EIGENFUNCTION.

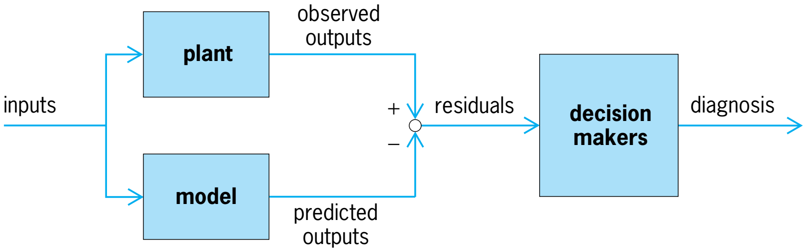


Fig. 1. Two stages of model-based fault detection and isolation. (*After R. Isermann and B. Freyermuth, eds., Proceedings of the IFAC SAFEPROCESS Symposium, Baden-Baden, Germany, September 10–13, 1991, Pergamon, 1992*)

Example of fault-tree analysis. The schematic of a simple electrical circuit in which a light is operated by a pair of three-way switches is shown in **Fig. 2**. (Such circuits are used in long hallways.) **Figure 3** shows the detailed fault tree of the circuit. The tree goes down to subcomponents (contacts of the switches) in order to illustrate more complex logic relations on this simple system. Note that nonfailure events (operating conditions) are also among the basic events because such conditions (the position of each switch) determine whether a particular failure event triggers the top event.

Example of consistency checking. Traditionally, a few fundamental variables, such as coolant temperature, oil pressure, and battery voltage, have been monitored in automobile engines by using limit sensors. With the introduction of onboard microcomputers, the scope and number of variables that can be considered have been extended. Active functional testing may be applied to at least one actuator, typically the exhaust-gas recirculation valve. Model-based schemes to cover the components affecting the vehicle's emission control system are gradually introduced by manufacturers. One approach (**Fig. 4**) uses analytical redundancy to monitor two groups of actuators (fuel injectors and exhaust gas recirculation) and four sensors (throttle position, manifold pressure, engine speed, and exhaust oxygen). By the appropriate selection of the model relations, the residuals are insensitive to the load torque and the vehicle's mass. The structured residual technique is used to support fault isolation. The critical issue

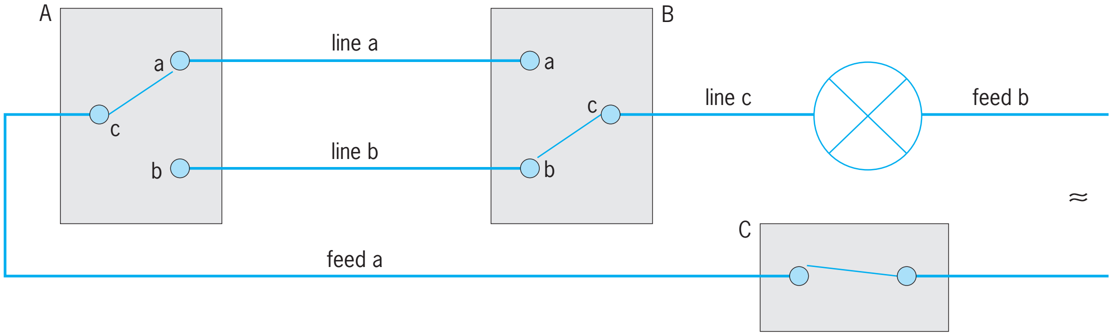


Fig. 2. Simple electrical circuit: a lamp operated by a pair of three-way switches.

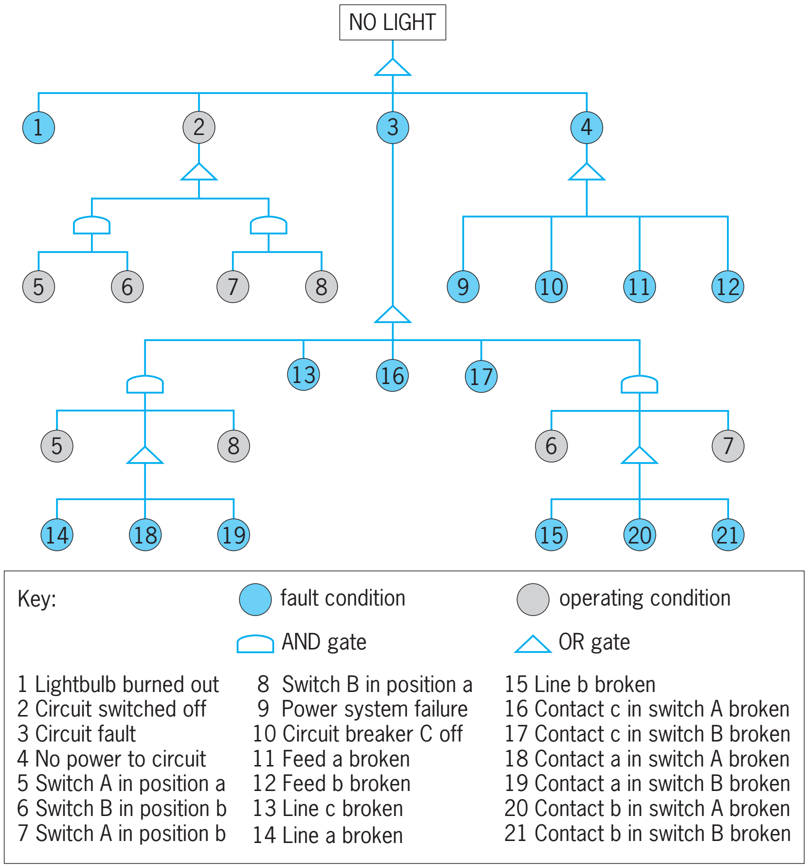


Fig. 3. Detailed fault tree of the circuit shown in Fig. 2.

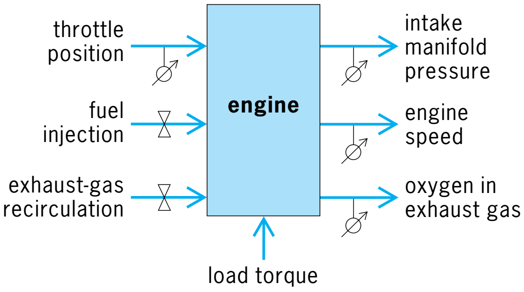


Fig. 4. Car engine system with onboard fault detection and diagnosis. (*After R. Isermann and B. Freyermuth, eds., Proceedings of the IFAC SAFEPROCESS Symposium, Baden-Baden, Germany, September 10–13, 1991, Pergamon, 1992*)

is to find sufficiently general models so that a single scheme may function well across an entire automobile product line and under widely varying operating conditions. *See* AUTOMOTIVE ENGINE; MICROCOMPUTER; MICROPROCESSOR. Janos J. Gertler

Bibliography. L. H. Chiang, R. D. Braatz, and E. Russel, *Fault Detection and Diagnosis in Industrial Systems*, Springer, 2001; P. L. Clemens, *Fault Tree Analysis*, 4th ed., Jacobs Sverdrup, 2002; J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*, 1998; J. Gertler, Survey of model-based failure detection and isolation in complex plants, *IEEE Control Sys. Mag.*, 8(7):3–11, 1988; R. Patton, P. Frank, and R. Clark (eds.), *Fault Diagnosis in Dynamic Systems*, 1989.

Fault and fault structures

Products of fracturing and differential movements along fractures in continental and oceanic crustal rocks. Faults range in length and magnitude of displacement from small structures visible in hand specimens, displaying offsets of a centimeter (1 cm = 0.4 in.) or less, to long, continuous crustal breaks, extending hundreds of kilometers (1 km = 0.6 mi) in length and accommodating displacements of tens or hundreds of kilometers. Faults exist in deformed rocks at the microscopic scale, but these are generally ignored or go unrecognized in most geological studies. Alternatively, where microfaults systematically pervade rock bodies as sets of very closely spaced subparallel, planar fractures, they are recognized and interpreted as a type of cleavage which permitted flow of the rock body. Fractures along which there is no visible displacement are known as joints. These include shear joints, formed by fracturing and imperceptible movement of the walls of the fractures parallel to fracture surfaces, and tension joints, formed by negligible or barely visible displacement of the walls of the fractures perpendicular to the fracture surfaces. Large fractures which have accommodated major dilational openings (a meter or more) perpendicular to the fracture surfaces are known as fissures. Formation of fissures is restricted to near-surface conditions, for example, in areas of crustal stretching of subsidence. When faulting takes place under conditions of high temperature or pressure, zones of penetrative shear flow may develop which are best described as ductile fault zones.

Locating faults. The recognition of faults in continental regions of moderate to excellent rock exposure is generally straightforward. Classically, the process of systematic geological mapping has proved to be a powerful method for locating faults. In essence, faults can be identified and tracked by recognizing in mapped patterns the truncation and offset of one or several bedrock units. Depending on the nature of the exposure and movement on the fault, truncation and offset might produce a simple horizontal shifting of the dominant mapped pattern of bedrock units; alternatively, the faulting might lead to a pattern of repetition or omission of specific rock formations within the geologic column.

Valuable physical signatures which reveal the presence of faults include abnormally straight topographic lineaments or fault-line scarps; aligned springs issuing from fractured and favorably displaced bedrock; intensely fractured rocks, perhaps with zones of angular chunks of brecciated, rotated materials; fracture surfaces naturally polished through the movement process and etched with striations or grooves; dragging (folding) of rock layers out of their normal orientation; loosely consolidated, ground-up rock flour or paste, commonly referred to as gouge; radically crushed, cataclastically deformed rocks known as mylonites; and alteration, silicification, or mineralization brought about by circulation of hot fluids through shattered bedrock. Faults are simple to locate in areas of active or very recent mountain building, especially where faulting has broken the ground surface and produced scarps. *See* MYLONITE.

In continental areas of very poor rock exposure, in the subsurface, and in ocean basins, faults are much more difficult and costly to locate. However, major faults are routinely discovered through application of geophysical methods, especially seismic, gravity, and magnetic surveying. Abrupt contrasts in the geophysical signatures of rocks at depth signal the sharp truncation of bedrock by faults and allow the pattern of faults to be mapped. The geophysical exploration of fault and fault structures in the ocean floor has completely changed the way in which geologists view the Earth and earth dynamics. It is known that major fracture zones, unlike any recognized in continental regions, exist in the ocean floor. These fractures, hundreds to thousands of kilometers in length and spaced at tens of kilometers, pervade the Mid-Oceanic Ridge system and are usually oriented perpendicular to the crest of the ridge segment which they occupy. Interpreting these to be enormous faults which accommodate the movement of newborn oceanic crust as it spreads bilaterally from the Mid-Oceanic Ridge system, J. T. Wilson named them transform faults. *See* GEOPHYSICAL EXPLORATION; MID-OCEANIC RIDGE; MARINE GEOLOGY.

Transform faults. Three fundamental types of transform faults exist: the first connects one ridge segment to another; the second connects a ridge segment (where new oceanic crust forms) to a trench (where oceanic crust is consumed through subduction); the third connects two trenches. Ridge-ridge transform faults, perhaps the easiest to visualize (**Fig. 1**), link parallel but offset ridge segments, and are characterized by shallow earthquake activity restricted to the part of the transform between the two ocean-ridge segments. The sense of movement along each ridge-ridge transform fault, as deduced from fault-plane solutions, is opposite to that which would explain the offset of the oceanic ridge. In essence, the movement along the transform fault does not produce the offset, but rather records the differential sliding-past (shearing) of the new ocean floor moving in opposite directions from the two ocean-ridge segments which are connected by the transform. Thus the transform faults serve, in this example, as part of the divergent plate boundary which accommodates sea-floor spreading and the movement of the two plates can be described by an imaginary pole of rotation located on the Earth's surface at a position defined by the intersection of great circles drawn perpendicular to points on the array of transform faults. The rate of relative movement along the ridge-ridge transform is a function of the rate of generation of new oceanic crust along the ridge and is generally on the order of several centimeters per year. *See* MOUNTAIN SYSTEMS; PLATE TECTONICS; TRANSFORM FAULT.

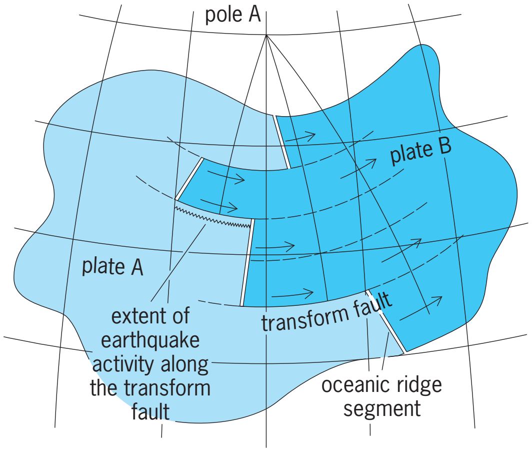


Fig. 1. Transform faults linking oceanic ridge segments. (*After W. J. Morgan, Rises, trenches, great faults, and crustal blocks, J. Geophys. Res., 73:1959–1982, 1968*)

Transform fault patterns are quite complicated, as is their evolution through time. Boundary zones between three adjacent tectonic plates (or triple junctions) are particularly difficult to evaluate with respect to geometry and kinematics (motions). Nonetheless, plate-tectonic analysis, including detailed assessment of transform faults, has had a revolutionary impact on understanding of the Cenozoic tectonic evolution of the Earth's crust. For example, scientists generally accept Tanya Atwater's hypothesis that the infamous San Andreas fault system of California is a transform fault boundary separating two enormous crustal plates, the North American and the Pacific. Relative movement between these plates is horizontal and right-handed, and in magnitude amounts to hundreds of kilometers.

Movements. This interpretation of the San Andreas Fault clearly demonstrates that some major fault systems in continental regions can be better understood in the context of the "new rules" afforded by study of transform faulting and plate tectonics. However, most fault systems in continental regions cannot be clearly and quantitatively linked to specific plate-tectonic movements or configurations, present or past, and they are analyzed and understood in an entirely different way; the guidelines for analysis are well established. In addition to describing the physical and geometric nature of faults and interpreting time of formation, it has been found to be especially important to determine the orientations of minor fault structures (such as striae and drag folds) which record the sense of relative movement.

Evaluating the movement of faulting can be difficult, for the apparent relative movement (separation) of fault blocks as seen in map or outcrop may bear little or no relation to the actual relative movement (slip). The slip of the fault is the actual relative movement between two points or two markers in the rock that were coincident before faulting

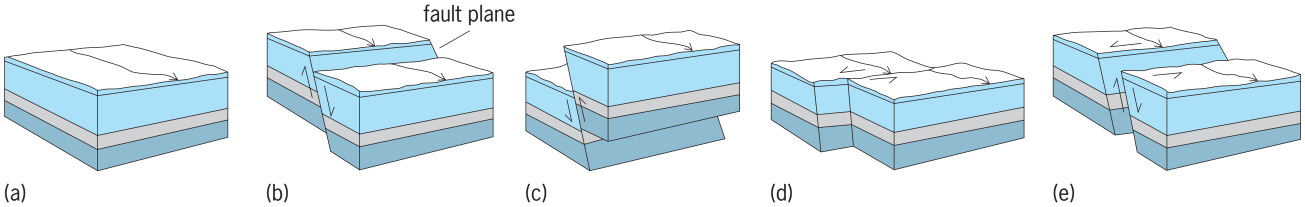


Fig. 2. Slip on faults. (*a*) Block before faulting; (*b*) normal-slip; (*c*) reverse-slip; (*d*) strike-slip; (*e*) oblique-slip. (*After F. Press and R. Siever*, *Earth*, *2d ed*., *W. H. Freeman*, *1978*)

(**Fig. 2**). Strike-slip faults have resulted in horizontal movements between adjacent blocks; dip-slip faults are marked by translations directly up or down the dip of the fault surface; in oblique-slip faults the path of actual relative movement is inclined somewhere between horizontal and dip slip. Strike-slip faults are described as left- or right-handed, depending on the sense of actual relative movement; dip-slip faults are described as normal-slip, thrust-slip, or reverse-slip, depending on the sense of actual relative movement and on the dip and dip direction of the fault surface. Listric normal-slip faults are a type of normal-slip fault in which the inclination of the fault decreases. Movement on such a curved fault surface produces a profound backward rotation of the upper block.

Recognizing even the simplest translational fault movements in nature is often enormously difficult because of complicated and deceptive patterns created by the interference of structure and topography, and by the absence of specific fault structures which define the slip path (**Fig. 3**). While mapping, the geologist mainly documents apparent relative movement (separation) along a fault, based on what is observed in plan-view or cross-sectional exposures. In the separation sense, left- and right-lateral faults are those displaying apparent horizontal shifts of rock in map view. Such shifts are said to be apparent because the left- or right-lateral offset might actually have been produced by dip-slip, not strike-slip, movements. Normal, reverse, and thrust are separation terms for faults, and again the usage of each is based on apparent offset in cross-sectional view and on the dip and dip direction of the fault.

Movement and offset on large, regional fault systems must be evaluated on the basis of displaced geologic terrains and abnormal stratigraphic relations. For example, 435 to 500 mi (700 to 800 km) of left-slip fault movement has been postulated in northern Mexico during the time period 150,000,000 to 170,000,000 years ago. The basis of the interpretation is truncation and offset of terrains of Precambrian and Paleozoic rocks in California and Arizona. Low-angle thrust movements in western Utah produced in Cretaceous time a 44-mi (70-km) west-to-east transport of thick Precambrian through Paleozoic miogeoclinal strata onto thin shelf and platform strata of the same age.

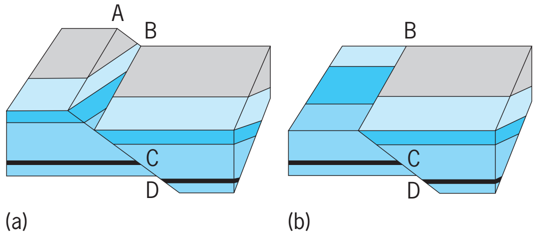


Fig. 3. Slip versus separation. (*a*) AB is the dip-slip. (*b*) After erosion of top of footwall block. The block has undergone a right separation. (*After M. P. Billings*, *Structural Geology*, *3d ed*., *Prentice-Hall*, *1972*)

Stress conditions. Theory on faulting is based on applied physics and engineering, and focuses on the stress conditions under which rocks break. The theory is almost exclusively concerned with the brittle behavior of crustal rocks; as such, it is most applicable to faulting at upper crustal layers in the Earth. Results of deformational experiments under controlled temperature, pressure, and strain-rate conditions bear importantly on modern understanding of the dynamics of faulting. The results of theoretical and experimental work provide insight into why faults can be conveniently separated into categories of normal-slip, thrust-slip, and strike-slip.

Forces which act on a rock body may be resolved by vector analysis into components of force acting in specific directions. These, in turn, can be converted to magnitudes and directions of stresses which tend to deform the body. This is done by dividing the force component by the surface area (of the body) on which it acts. Two types of stresses are distinguished, normal stress (σ), which acts perpendicular to a given surface, and shear stress (τ), which acts parallel to the given surface. Stress analysis, in effect, evaluates the magnitude of shear and normal stresses acting in all directions throughout a body and predicts the orientations of the surfaces along which faulting should occur.

Evaluating the distribution of stresses in a body that is acted upon by forces discloses that there are three unique directions within the stress field, called principal stress directions. These stress directions are mutually perpendicular, and the value of shear stress equals zero only along these three directions. Furthermore, one of the principal stress directions is characterized by the maximum value of normal stress (σ_1) within the system, and another (σ_3) is characterized by the minimum normal stress. Maximum shear stress values ($\tau_{\max}$) occur along lines oriented 45° to the principal stress directions.

Two-dimensional mathematical analysis demonstrates that the magnitude of normal stress (σ) on any plane (that is, on any potential surface of faulting)

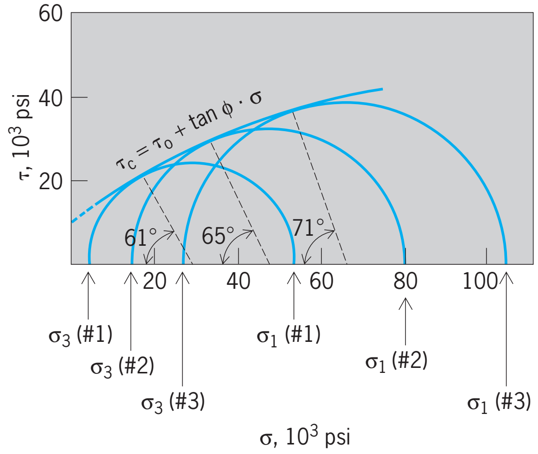


Fig. 4. Mohr circle diagram constructed from three sets of compressional tests of limestone. Angles (61°, 65°, and 71°) are 2θ values. 1 psi = 6.89 kPa.

in the body is given by Eq. (1), where σ = greatest

$$\sigma = \frac{\sigma_1 + \sigma_3}{2} - \frac{\sigma_1 - \sigma_3}{2} \cos 2\theta \qquad (1)$$

principal normal stress, σ_3 = least principal normal stress and θ = angle between the greatest principal stress axis and the direction of shearing stress (in the plane for which σ is being evaluated). The magnitude of shear stress (τ) on that same plane is given by Eq. (2). The distribution of paired values of normal

$$\tau = \frac{\sigma_1 - \sigma_3}{2} \sin 2\theta \qquad (2)$$

stress and shear stress as a function of θ is such that shear stress (τ) is zero at values of 0°, 90°, 180°, and 270°, but attains maximum values at 45°, 135°, 225°, and 310°. Normal stress reaches a maximum at 90° and 270°, but is minimal at 0° and 180°. These variations in normal and shear stress values may be portrayed on a Mohr circle diagram, a graphical representation of the above equations (**Fig. 4**). Points on the periphery of the circle have coordinates (σ,τ) which correspond in value to normal stress and shear stress on a plane which makes an angle of θ with σ_1, the greatest principal stress direction.

Given such a stress distribution, and assuming that the differential stress conditions ($\sigma_1 - \sigma_3$) exceed the strength of the rock body, the orientation of faulting can be determined. The Mohr-Coulomb law of failure [Eq. (3)] predicts that faulting should occur

$$\tau_c = \tau_0 + \sigma \tan\phi \qquad (3)$$

at a critical shear stress level (τ_c) where τ_0 = cohesive strength of the rock, σ = normal stress on the fault plane, and ϕ = angle of internal friction. The coefficient of internal friction, tan ϕ, equals σ/τ at failure. For most rocks, the coefficient of internal friction has a value between 0.4 and 0.7; thus the angle of internal friction, ϕ, commonly varies from 20 to 35°. The value of θ for such internal friction typically ranges from 27 to 35°, and is often 30°. In practice, the failure points on Mohr circle diagrams, as generated through experimental deformation, do not conform to the ideal straight-line failure envelope predicted on the basis of the Mohr-Coulomb law. Rather, the failure envelopes are smoothly curved in a way that describes an increase in 2θ with an increase in confining pressure (σ_3). Predictions as to when and at what angle faulting should occur are further complicated by variables of temperature, strain rate, pore-fluid pressure, and presence of fractures. In fact, fundamental questions have been raised regarding the extent to which theory and short-term experimental work can be applied to some natural geological systems.

What arises from theory and experiments is that fractures form in an orientation such that they contain the axis of intermediate stress (σ_2) and make an angle of θ (commonly around 30°) with σ_1 and 90 − θ with σ_3. Since the Earth's surface has no shear stress, principal stress directions near the Earth's surface tend to be vertical and horizontal, and depending on the relative configuration of the principal stresses, will give rise to thrust-slip, normal-slip, or strike-slip faults. The direction of movement on these faults is such that the wedge receiving the greatest compressive stress moves inward, whereas the wedge receiving the least compressive stress moves outward.

Examples. There are many excellent natural examples of normal-slip, thrust-slip, and strike-slip faults. The Basin and Range Province of the western United States displays a unique physiographic basin/range style because of pervasive large-scale normal-slip faulting (**Fig. 5**), which resulted from regional crustal extension within the last 15,000,000 years. Normal-slip faults accommodate extension.

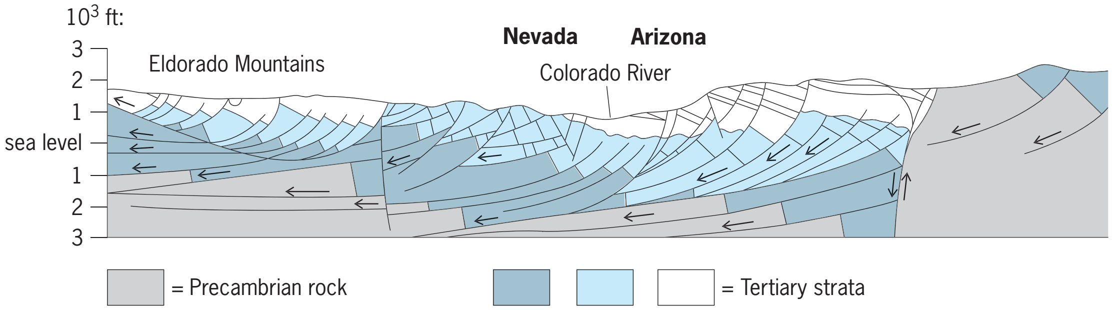


Fig. 5. Cross section of listric normal faults in the Basin and Range Province. (*After R. E. Anderson*, *Geologic Map of the Black Canyon 15-Minute Quadrangle, Mohave County*, *Arizona*, *and Clark County*, *Nevada*, *USGS Map GQ-1394*, *1978*)

Thrust-slip faults are an integral part of the tectonic framework of the southern Appalachian Mountains, the Sevier orogenic belt of western Utah and western Wyoming/eastern Idaho, and the Canadian Rockies. In mountain belts throughout the world, thrusting has played a major role in accommodating crustal shortening during mountain building. Major mechanical questions have arisen as to how enormous masses of crustal rocks can be thrust tens or hundreds of kilometers, since the apparent force required should have crushed the rock mass before it moved. The paradox of regional overthrusting has led to theories of faulting that emphasize the importance of factors such as high pore-fluid pressure, gravitational sliding, viscous creep of ductile materials, and underthrusting.

The best-documented strike-slip faulting is concentrated near margins of lithospheric plates, but such faulting has occurred in foreland tectonic regions as well, including the Rocky Mountains of the American West and mainland China. A major theme that has emerged from the study of high-angle faults, especially strike-slip faults, is that ancient faults in basement rocks are commonly reactivated in post-Precambrian time, producing zones of concentrated, superposed strain. This interpretation has been used to explain the classic monoclinal uplifts of the Colorado Plateau. *See* GRABEN; HORST; STRUCTURAL GEOLOGY.

George H. Davis

Bibliography. M. P. Billings, *Structural Geology*, 4th ed., 1986; J. G. Dennis, *Structural Geology: An Introduction*, 1987; R. E. Powell and J. C. Matti (eds.), *San Andreas Fault System: Displacement, Palinspastic Reconstruction, and Geologic Evolution*, 1993; F. Press and R. Siever, *Understanding Earth*, 3d ed., 2000; R. J. Twiss and E. M. Moores, *Structural Geology*, 1992; J. T. Wilson, A new class of faults and their bearing on continental drift, *Nature*, 207:343–347, 1965.

Fault-tolerant systems

Systems, predominantly computing systems, telecommunication systems, and other computer-based systems, which tolerate undesired changes in their internal structure or external environment. Such changes, generally referred to as faults, may occur at various times during the evolution of a system, beginning with its specification and proceeding through its utilization. Faults that occur during specification, design, implementation, or modification are called design faults; faults that occur during utilization are referred to as operational faults. Design faults are due to mistakes made by humans or by automated design tools in the process of specifying, designing, implementing, or modifying a system. Operational faults can be physical or human and can occur internally or externally. Examples include component failures (physical-internal), temperature and radiation stress (physical-external), mistakes by human operators who are integral parts of the system (human-internal), and inadvertent or deliberate interference by humans who interact with the system (human-external).

Dependability and performability. Both design faults and operational faults can affect a system's ability to serve its user(s). With respect to a specification of expected service, such ability is referred to as dependability (the ability of a system to deliver the specified service). System service is classified as proper if it is delivered as specified; otherwise it is improper. System failure is identified with a transition from proper to improper service and is a consequence of errors caused by faults. Dependability is quantified by reliability, which is a measure of continuous delivery of proper service, and availability, which quantifies the alternation between deliveries of proper and improper service. More refined assessments of a system's ability to serve are provided by measures of performability (the probability distribution of how well a system performs over a specified period of time). In particular, such measures can account for degraded levels of service that, according to failure criteria, remains proper.

Specific requirements for dependability and performability can vary dramatically according to the nature of the service provided. At one extreme, degraded performance or failure may simply inconvenience the user; in this case, the requirements are relatively modest. At the other extreme, a system failure may result in the loss of human lives; here, continuous proper service is essential and, hence, ultrahigh reliability becomes the dominating requirement. *See* RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

Fault prevention and fault tolerance. To satisfy a specified set of dependability and performability requirements, faults can be dealt with by prevention or tolerance. Fault prevention is the more traditional method. It attempts to ensure a fault-free system through methods of fault avoidance, which eliminate faults—for example, design methodologies, quality control methods, or radiation shielding; or methods of fault removal, which find and remove faults prior to system utilization—for example, testing and verification. Fault tolerance techniques are based on the premise that a complex system, no matter how carefully designed and validated, is likely to contain residual design faults and is likely to encounter unpreventable operational faults.

To describe how systems tolerate faults, it is important to distinguish the concepts of fault (defined above) and error. Relative to a description of desired system behavior at some specified level of abstraction, an error is a deviation from desired behavior caused by a fault. Since desired behavior may be described at different levels of abstraction—for example, the behavior of computer hardware may be described at the circuit level, logic level, or register-transfer level—a variety of errors can be associated with a given fault. Moreover, a fault can cause an error at one level without causing an error at another level. A fault is latent until it causes an error, and it is possible for a fault to remain latent throughout the lifetime of a system, thus never causing an error.

The distinction between fault and error is particularly important in the case of design faults. For example, a software design fault (commonly referred to as a bug) may remain latent for years before the software is finally executed in a manner that permits the bug to cause an execution error. The sense in which a system tolerates a fault is typically defined as some form of restriction on the errors caused by that fault. Often, though not always, this term refers to the prevention of errors at a level that represents the user's view of proper system behavior. Errors at this level coincide with system failures and, accordingly, a fault is tolerated if it does not cause a system failure. A system is fault-tolerant if all faults in a specified class of faults are tolerated.

Fault tolerance techniques. A variety of fault tolerance techniques are employed in the design of fault-tolerant systems. Generally, such techniques attempt to prevent lower-level errors (caused by faults) from causing system failures. Using various types of temporal, structural, and informational redundancy, such techniques either mask a fault (no errors are propagated to the faulty subsystem's output) or detect a fault (via an error) and then effect a recovery process which, if successful, prevents system failure. In the case of permanent internal faults, the recovery process usually includes some form of reconfiguration (replacement of a faulty subsystem with a spare; use of an alternate program) which prevents the fault from causing further errors. Many of these methods invoke multiple computations whose results are used to mask faults, via a decision mechanism that produces a single result, or to detect faults through discrimination of differences among their values. This assumes that, with high probability, faults do not cause similar errors among multiple results. Hence, in the case of design faults, simple replication computation is precluded; instead, multiple computations need to be realized by diverse designs which minimize the probability of similar errors.

Typically, a fault-tolerant system design incorporates a combination of fault tolerance techniques which complement the techniques used for fault prevention. The choice of these techniques and, more generally, the specific nature of the system's hardware and software architecture is highly dependent on both types of faults anticipated and the dependability and performability requirements. *See* COMPUTER ARCHITECTURE; INFORMATION THEORY; SOFTWARE ENGINEERING. John F. Meyer

Bibliography. S.-T. Levi and A. K. Agrawala, *Fault Tolerant System Design*, 1994; J. F. Meyer et al. (eds.), *Dependable Computing for Critical Applications, Two*, 1992; D. K. Pradhar, *Fault-Tolerant Computing: Theory and Techniques*, vol. 1, 2d ed., 1993; D. P. Siewiorek and R. S. Swarz, *Reliable Computer Systems: Design and Evaluation*, 3d ed., 1998.

Feather

A specialized keratinous outgrowth of the skin, which is a unique characteristic of birds. Feathers are highly complex structures that provide insulation, protection against mechanical damage, protective coloration, and also function significantly in behavior. One special functional role is in flight, where feathers provide propulsive surfaces and a body surface aerodynamically suitable for flight. Feathers are used in maintenance of balance and occasionally in the capture of prey and various specialized displays.

Distribution. The number of feathers on birds is determined by a variety of parameters but, in general, large birds have more feathers than small birds. Feather counts vary from 940 on the ruby-throated hummingbird to about 4000 on sparrow-sized birds, 5000 on blackbirds, and more than 25,000 on a whistling swan. There is a significant correlation between body weight and the relative weight of the feathers, so that large birds have proportionately more plumage per surface area than small birds. No significant taxonomic correlation is found in feather number or weight; presumably the relative feather mass is involved in the conservation of metabolic heat. Thus the feather mass of various species is more likely to be influenced by ecological and physiological factors than by taxonomic position.

The uniform distribution of feathers, which appear to cover a bird almost completely, is only superficial. Definitive (as opposed to down) feathers grow from restricted tracts (pterylae) and are absent in the intervening areas (apteria). Number of feathers often varies according to body part, and the head and neck often carry 40–50% of the total number. The pterylae have well-defined patterns which are at least partially specific for different taxonomic groups. The study of these patterns is termed pterylography.

Structure. A representative definitive feather (**Fig. 1**) contains a single long central axis which supports a row of small branchlike structures along each side (barbs). The central structure consists of a hollow, vaneless, cylindrical basal section, called the calamus, and a solid, angular portion, the rachis. The calamus has openings at both ends (the upper and lower umbilicus), the function of which is only incompletely known. The umbilici and calamus are especially well developed in powder down, and the umbilici may be related to the growth of these feathers. Barbs form the vane, or web, of the feather. Individual barbs branch off the rachis at variable angles and point toward the outer tip of the feather.

The barbules are small branches from the barbs. They lie in the same plane as the barbs and arise in rows from their anterior and posterior surfaces. The anterior barbules have a flattened base and a series of small hooklike projections called hamuli; the hamuli attach to the proximal ridge of the posterior barbules of the next barb, forming an interlocking structure characterized by its great strength and light weight (**Fig. 2**). All feather types consist basically of these structural elements. The distribution

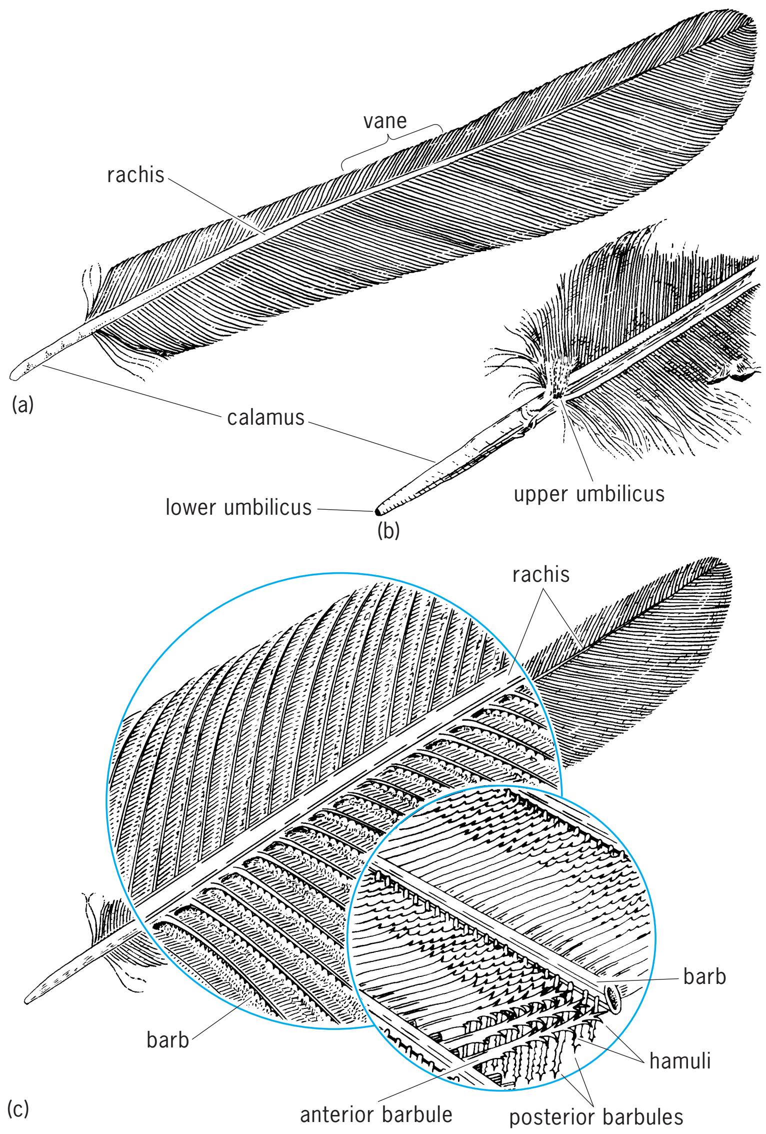


Fig. 1. Typical flight feather. (*a*) Full view. (*b*) Detail of base. (*c*) Detail of vane. (*After J. Van Tyne and A. J. Berger*, *Fundamentals of Ornithology*, *2d ed.*, *John Wiley and Sons*, *1976*)

and interrelationships of these elements produce the three-dimensional, macroscopic appearance of the feather.

Specializations. Most of the superficial feathers are contour feathers (pennae). These include the large flight feathers (remiges) of the wing and the long tail feathers (rectrices). Other common feather types include the down feathers (plumulae), intermediate types (semiplumes), and filoplumes (**Fig. 3**). Pennae grow only in pterylae and form the major contours of the body. Semiplumes lack hamuli and thus lack a firm vane. They occur at the periphery of pterylae, integrate structurally with pennae, and serve to provide insulation, bouyancy, and increased flexibility at the base of moving parts. Filoplumes are specialized structures and may be decorative or sensory or both. They are always associated with the papillae of contour feathers. Filoplumes have only a reduced vane at their tip or lack a vane completely. The reduced vane (vibrissa) is hairlike in appearance and presumably has a tactile function. Vibrissae are well developed in the eyelashes of several species and in the stiff (rictal) bristles found at the base of the bill. Down feathers on adults are usually concealed by the contour feathers. On hatchlings of most species the entire feather complement consists of downy feathers. The degree of nestling feathers (neossoptiles) ranges from sparse to dense and, in many cases, is related to chick behavior and the ecology of the nest. Most natal downs lack a rachis and are not confined to pterylae. Their primary function is insulative. Powder down, found only in certain taxonomic groups, are continually growing feathers. They produce a soft, friable material used in the cleaning and maintenance of other feathers.

Naked, or featherless, areas of adult birds serve several functions. For example, combs, wattles, bare eye rings, casques, and other structures on or near the head are used in sexual and mating displays; parts of the legs and feet may be used in the regulation of body heat, or for the protection, by shading, of the eggs and young; the naked head and neck in carrion feeders may serve a sanitary function.

Cytology. Adult feathers are essentially dead structures. However, developing feathers consist of three rather distinct cell types. These form the cuticle, cortex, and medulla, the morphologically distinct elements of the feather. The proportions of these elements vary within any particular type of feather. In the medulla, especially, the relationship between inter- and intracellular spaces may vary and influence the appearance of the feather, particularly its color. Cells of the developing cortex show high cytoplasmic ribonucleic acid (RNA) activity, as occurs in cells undergoing rapid proliferation and protein synthesis. *See* RIBONUCLEIC ACID (RNA).

Chemistry and molecular structure. Feathers are ectodermally derived proteinaceous material. The main protein, β-keratin, has a molecular weight of approximately 11,000 and is characterized by its general insolubility, a result of the large numbers of intramolecular disulfide bonds. Although the amino acid composition of feather parts may differ, serine, glycine, cystine, and proline are the most frequent residues. *See* PROTEIN.

The solubilized proteins from feathers are highly heterogeneous upon separation by electrophoresis.

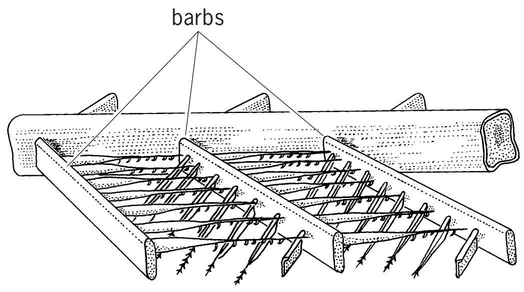


Fig. 2. Parts of three barbs from a vaned feather, showing the hooking mechanism of the overlapping barbules. (*After J. C. Welty*, *The Life of Birds*, *Saunders*, *1962*)

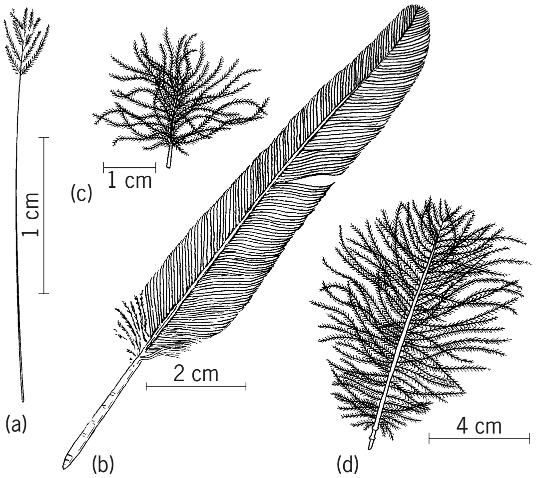


Fig. 3. **Types of feathers, (*a*) Filoplume. (*b*) Vane or contour. (c) Down. (*d*) Semiplume. (*After J. C. Welty, The Life of Birds, Saunders, 1962*)**

Combined with differences in physiochemical and immunological properties, the complexity indicates the presence of a large number of monomeric, individual gene products. This interpretation is supported by the existence of species-specific keratin polypeptides and experiments with messenger RNA (mRNA) and complementary deoxyribonucleic acid (DNA) which indicate the presence of over 200 keratin genes in the total genome. Differences in molecular protein structure and distribution reflect the interaction of structural and regulatory gene products. In certain genetic mutants the abnormal gross feather structure is the result of the same interactions.

The complete amino acid sequence is known for keratin monomers from two species. Physical analysis shows each monomer, or gene product, to consist of a folded and an amorphous portion. Sequence differences appear limited to the latter. The folded portions interact between units to form larger, highly organized, molecular assemblies. The folded portion may also contain the immunogenetic site of the protein. Thus, structural conformation is related to an evolutionary conservativeness. The individual polypeptides form small microfibrils. These, in turn, are organized into filaments which are cemented together in an amorphous matrix.

Development and molt. Feather development is a complex process and has been studied in detail only in the contour and flight feathers of domestic birds. The penna develops initially as a domelike elevation of the epidermis that has a dermal core. Each developing feather grows outward by a rapid proliferation of cells at its base (**Fig. 4**). A tapered epidermal cylinder is produced with a dermal core that contains vessels which provide nutrients to the developing structure. Subsequent feather development depends on accurate alignment, in both time and space, of rows of epidermal cells to form the rachis, barbs, and barbules. This is especially true in the highly structured adult contour and flight feathers, in which growth is adjusted for the subunits of the vane to hook together properly. In addition, specialized feather structures, such as the tips on the feathers of cedar waxwings and similar structures in other species, are produced by modification of the same basic mechanism.

Histogenesis. As development proceeds, three epidermal areas can be distinguished. The outer layer, which forms a sheath; a thick intermediate layer, representing the primordia of the rachis and barb, which later becomes keratinized; and an inner layer, which forms an envelope about the pulp. All three layers are derived from a collarlike ring of embryonic epidermal cells. In all cases a remnant of the original dermal cells, surrounded by epidermal cells, remains at the papilla. This is necessary for formation and production of subsequent generations of feathers. In newly hatched chicks the epidermal sheath ruptures as it dries, allowing feather parts to unfurl. As outward growth proceeds, the pulp is simultaneously withdrawn toward the feather base, producing a completely keratinized, nonliving structure.

Endocrine regulation. Many aspects of feather growth are dependent on hormones. Evidence exists which suggests that thyroid hormones are immediately responsible for molt by acting directly on the feather papilla, but there is a great deal of interspecific difference in the quality and timing of response to thyroxine administration. Pigmentation is, in varying degrees, determined by thyroid, sex, and gonadotrophic hormones. Hormone imbalances, especially in reference to thyroid secretion, can cause changes in the pattern and melanin content of feathers, as well as affect the timing of the molt.

Periodicity of molt. Feathers normally undergo attrition because of the physical abuse attendant to the normal activity of birds. In most species, feathers are replaced completely at least annually, and many of the feathers are replaced more frequently. The actual molt involves an exquisite integration of the sequence and relative timing of a series of events and

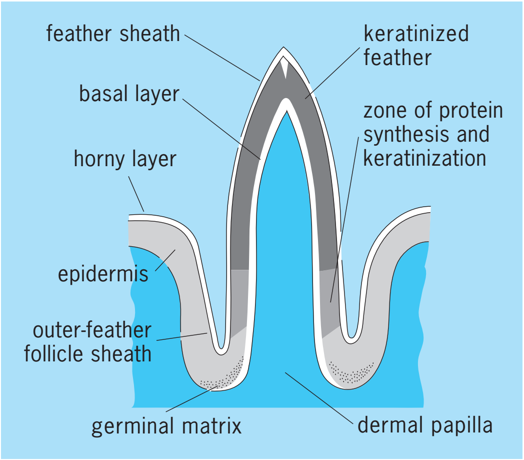


Fig. 4. **Section through growing feather follicle. (*After R. I. C. Spearman, The keratinization of epidermal scales, feathers, hairs, Biol. Rev., 41:59–69, 1966*)**

is clearly of adaptive value. Considerable energy is expended in molt, in the biosynthesis of new keratin, in the associated changes in thermal conductance, and in associated behavioral changes. For this reason molting is usually scheduled so as not to interfere with other energy-demanding phenomena of the annual cycle, most notably reproduction and migration.

Molt patterns. The sequence of feather molt is surprisingly orderly. Penguins, which shed large patches of feathers in an irregular pattern, are an exception. In most species the power of flight is retained during molt. The molt, that is, the normal shedding of feathers and their replacement by a new generation of feathers, is a single growth process which is actively concerned only with the production of the new generation of feathers. The old feathers are pushed out of the follicles passively.

Coloration. The patterns and coloration in feathers are controlled by mechanisms active during their formation. Obviously, hormonal and nutritional conditions are important during this period. Less well understood are the systems which produce and control the repetition and register so apparent in the plumage pattern. Such patterning involves the control of follicle spacing, feather angle, and the initiation and termination of the pigmentation process in individual feathers. Superficial patterns of this type provide the basis for species recognization.

There are basically two types of feather colors: schemochromes, which are colors that originate within the feather structures; and biochromes, which are colors that are caused by the presence of true pigmentation.

Schemochromes. The schemochromes are derived from morphological variations in the rachis, barbs, and barbules. Pigments are not necessary for the production of structural colors, but they are generally present and modify the colors considerably. The simplest case of structural coloration is whiteness, which is the result of almost complete reflection and refraction of incident light by the minute, colorless, transparent feather surfaces, interstices, and air-filled spaces. Noniridescent blues, such as in the blue jay, are another common structural color produced when a fine, colorless framework of barbs overlies a dark layer which contains melanin granules. The colorless layer reflects the blue wavelength, while other wavelengths pass through this layer and are absorbed by the pigment below. Other noniridescent colors, such as various greens or olive, are also associated with the presence of pigments in the absorbing layer. These include both melanins and carotenoids. *See* TYNDALL EFFECT.

Iridescent feathers, such as are found on hummingbirds, sunbirds, and the rock dove, are produced independently of pigmentation by the interference of light by a thinly laminated structure in the barbules. Barbs of iridescent feathers are invariably broad and flat with blunted ends, distinctly pigmented, and generally lacking barbules. These broad surfaces often overlap, and the feather surface then acts as a diffraction grating. Colors are caused by the interference of light waves reflected from the surfaces of extremely thin, colorless laminae or film-like plates which compose the outer surfaces of the barbules. The various colors are produced by differences in the angle of the surfaces and changes in the angle of light incident to the surface.

Biochromes. The most common biochromes are melanins. Melanins are synthesized from the amino acid tyrosine in specialized cells (melanophores) and are deposited as discrete particles or granules. There are probably no consistent chemical or structural differences between the eumelanins and phaemelanins even though they produce slightly different colors. Melanins are responsible for a wide range of colors, from dull yellow through red-brown to dark brown and black. *See* PIGMENTATION.

Carotenoids. The most spectacular colors in feathers are produced by the carotenoid pigments. Carotenoids are found mainly in the apical portion of the barbs, but in intensely pigmented feathers the entire barb may be pigmented. In certain cases intense pigmentation is accompanied by morphological changes in the feather, such as flattening of the barb in one plane and the absence of barbules. This provides a better exposure of the pigment. Early workers reported the absence of carotenoids in the feathers of young birds, but more evidence for such statements is necessary. *See* CAROTENOID.

As birds cannot synthesize carotenoids, they depend on their food for these pigments. Feather carotenoids are generally highly oxidized forms (commonly keto or hydroxy derivatives) of ingested pigments. Thus, birds obviously possess the metabolic capability of modifying these pigments. Accumulating evidence indicates that these pathways may be common to many groups of birds. Elucidation of the metabolic pathways for the production of feather carotenoids may be fruitful in understanding both the evolution of such pathways and the birds which possess them.

Color variability regulation. Many of the seasonal displays of feather pigments, especially in the males of migratory songbirds, depend on carotenoids. These deposits, as well as the pigments responsible for sexual dichromatism, presumably are controlled by hormonal mechanisms. In the brightly colored scarlet tanager, for example, the biochemical difference between the feathers of the males and the females, and between the colors of breeding and nonbreeding males, represents only a relatively small difference in the chemistry of the pigment molecule. This difference produces an enormous visual effect, however, and is of great significance to the biology of this species. The chemical differences in the feather pigments of the sexes and the seasonal differences in the males may be mediated by a single enzyme, determined by a single gene. Conceivably, the activity of this gene could be controlled by a single hormone. Additional studies which link pigment chemistry, diet, genetics, and natural history will further understanding in this area. Other genetic effects on pigments include the absence of pigments in normally pigmented individuals (albinism) or an overabundance of melanins (melanism) or carotenoids (xanthophyllism). *See* SEXUAL DIMORPHISM.

Porphyrin compounds. An unusual pigment has been isolated from members of the touraco family (Musophagidae). This pigment occurs in two forms in feathers, the red turacin and a green, oxidized derivative, turacoverdin. Chemically, this pigment is a copper-containing uroporphyrin. A related pigment, but one lacking copper, has been found in the bustards (Otididae). Free porphyrins (coproporphyrin) have also been reported in the down feathers of owls.

Adherent colors. Another type of coloration, produced by neither internal pigmentation nor structure, is the adherent colors. They are chemicals, commonly iron oxides, that are applied, presumably unintentionally, to feathers during preening.

Function. Many functional aspects of feathers are dependent on their posture and position. Feather posture is adjusted chiefly by bands of nonstriated muscle fibers in the dermis (m. arrectores plumarum). Typically each pennaceous feather follicle is connected by pairs of muscles to four adjacent follicles. The depressor and erector muscles in each pair provide fine control over the position of individual feathers. Receptors in the skin near the feather may sense pressure or vibration of the feather and thereby provide information for further adjustment of position.

Thermoregulation. A major physiological role of feathers is to provide insulation. This is accomplished by regulating the configuration of feather and skin in such a way that differing amounts of air are trapped in the dead space so formed. A second mechanism for control of heat dissipation is the balance of the exposure of feathered and unfeathered body parts. *See* THERMOREGULATION.

Waterproofing. Feathers act as a protective boundary in their role of providing waterproofing. Water repellency is a structural feature of feathers and is the result of precise geometric relationships between the diameter and spacing of barbs and barbules. Preening appears to be more important in the maintenance of this structure than it is for the application of oils or any other natural product, as was once thought.

Behavior. A third function of the surface configuration and overall pattern of feathers is in the area of behavioral adaptations. These may be of two types. First is concealment, when the bird is cryptically marked to match its background and escape detection. Other types of concealment include disruptive patterns, which break the familiar outline of the bird, and mimicry, a rare situation when one species resembles another. The second type consists of various types of advertisement. For example, unique combinations of feather color and pattern may label the species as inedible, serve to maintain contact among members of a flock in a gregarious species, or provide a means of attracting and stimulating members of the opposite sex. *See* PROTECTIVE COLORATION.

Evolution. Feathers presumably derived from scales or scalelike structures present on the reptilian ancestor of birds. Chemical and structural evidence indicates the presence of similar proteins in these structures. The form and process of derivation of the primitive feather are unknown. However, there is some agreement that primitive feathers were rather generalized in structure and arose in conjunction with the metabolic and thermoregulatory demands of avian precursors. This generalized feather, in turn, gave rise to the more specialized downy and contour plumages. *See* SCALE (ZOOLOGY). Alan H. Brush

Bibliography. A. H. Brush (ed.), *Chemical Zoology*, vol. 10, 1976; A. H. Brush (ed.), *Symposia on Feather Structure: Proceedings of the 16th International Ornithological Congress*, *Canberra*, *1974*, 1976; A. M. Lucas and P. B. Stettenheim, *Avian Anatomy*: *Integument*, pts. 1 and 2, USDA Handb. 362, 1972.

Feedback circuit

A circuit that returns a portion of the output signal of an electronic circuit or control system to the input of the circuit or system. When the signal returned (the feedback signal) is at the same phase as the input signal, the feedback is called positive or regenerative. When the feedback signal is of opposite phase to that of the input signal, the feedback is negative or degenerative.

The use of negative feedback in electronic circuits and automatic control systems produces changes in the characteristics of the system that improve the performance of the system. In electronic circuits, feedback is employed either to alter the shape of the frequency-response characteristics of an amplifier circuit and thereby produce more uniform amplification over a range of frequencies, or to produce conditions for oscillation in an oscillator circuit. It is also used because it stabilizes the gain of the system against changes in temperature, component replacement, and so on. Negative feedback also reduces nonlinear distortion. In automatic control systems, feedback is used to compare the actual output of a system with a desired output, the difference being used as the input signal to a controller. These two points of view will be considered separately in the following discussion. However, the analysis of both feedback amplifiers and electromechanical control systems can be made on a common basis; from the point of view of analysis the two have much in common.

Amplifier feedback. Feedback can be introduced into an amplifier by connecting to the input a fraction of the output signal. An amplifier will, in general, have better frequency-response characteristics when the system has feedback than when there is no feedback. The system can be designed to have a wider bandwidth and more nearly ideal frequency-response characteristics. Further, harmonic distortion due to nonlinear transistor or tube characteristics can be reduced by using feedback. *See* DISTORTION (ELECTRONIC CIRCUITS).

The use of feedback in an improperly designed system, however, can produce a system with worse characteristics. Amplifiers can become oscillators when feedback is used in an improperly designed system. Furthermore, feedback is not always intentional, and amplifiers, particularly radio-frequency amplifiers, are often unstable because of unintentional (parasitic) feedback paths in the circuit. For a discussion of amplifier frequency response and bandwidth *see* AMPLIFIER

A system with feedback can be shown in block diagram form (see **illus.**). The sinusoidal input signal is $\mathbf{E}_s$ and the amplifier gain is **A**, which is a function of frequency. When there is no feedback, $\mathbf{E}_o = \mathbf{AE}_s$ because $\mathbf{E}_i = \mathbf{E}_s$. When there is feedback, $\mathbf{E}_o = \mathbf{E}_s + \mathbf{E}_f$.

Since $\mathbf{E}_o = \mathbf{AE}_i$, and $\mathbf{E}_f = \beta\mathbf{E}_o$, the overall gain of the system with feedback is then shown by the equation below.

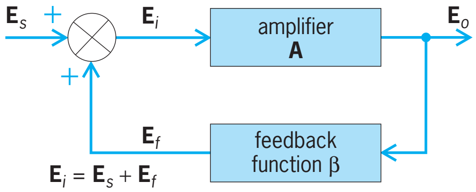
$$\frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{\mathbf{A}}{1 - \mathbf{A}\beta}$$

This formula for the gain of an amplifier with feedback indicates the effect of feedback upon the frequency response of the amplifier without feedback. The $(1 - \mathbf{A}\beta)$ term in the denominator is a complex number. Therefore, the magnitude and phase angle of the gain of the amplifier with feedback will differ from the gain of the amplifier without feedback. The amount of the difference depends upon the value of the **A** and the β terms, and no general statements can be made.

If there is a frequency for which $\mathbf{A}\beta = 1$, the denominator of the expression for the gain will be zero while the numerator will not be zero. When this occurs, the amplifier will oscillate at approximately the frequency for which $\mathbf{A}\beta = 1$. Furthermore, if there is a frequency for which the magnitude of $\mathbf{A}\beta$ is greater than unity and the phase angle of $\mathbf{A}\beta$ is 0°, 360°, or any integral multiple of 360°, the amplifier will oscillate.

Positive and negative feedback. The terms positive feedback and negative feedback are used to denote the type of feedback found in certain electronic circuits. *See* NEGATIVE-RESISTANCE CIRCUITS.

When the magnitude of the denominator of the feedback equation is greater than unity, the overall gain with this (negative) feedback will be less than the gain with no feedback, and stable operation will result. If the magnitude of $\mathbf{A}\beta$ is much greater than one, $\mathbf{A}/(1 - \mathbf{A}\beta)$ is approximately $-1/\beta$ and does not depend on **A**, thus minimizing the effect of unstable, inaccurate elements on the circuit. This result is the principal benefit of negative feedback.



Block diagram of feedback circuit.

If the magnitude of the denominator in the displayed equation is less than unity, then the overall gain with this positive feedback will be greater than the gain without feedback. Under these circumstances the circuit may be unstable and either oscillate or drive its output to the limit of the power-supply voltage level and remain there.

It follows that both the magnitude and the phase angle θ of the product $\mathbf{A}\beta$ are important in determining whether the feedback is positive.

The effect of feedback upon the frequency response of an amplifier can be determined from the expression $\mathbf{A}/(1 - \mathbf{A}\beta)$. The performance in a particular case depends upon the behavior of **A** and β as functions of frequency.

Analysis and design. There are several methods for analyzing and designing feedback amplifiers. In those cases where the **A** and β terms in the previous equation can be identified from the circuit, the possibilities of oscillation due to the feedback can be determined by plotting the amplitude and phase of $\mathbf{A}\beta$ over the frequency range from zero to a sufficiently high frequency. The examination is often made by plotting the magnitude of $\mathbf{A}\beta$ against the phase angle in polar coordinates. It may be done also by plotting the magnitude of $\mathbf{A}\beta$ in decibels and the phase of $\mathbf{A}\beta$ against the logarithm of the frequency for a sufficiently wide frequency range. The second method also indicates whether the amplifier will be satisfactory for audio applications.

There are many cases where the circuit is sufficiently complex that more powerful tools are required. The simple block diagram of the illustration is not applicable. Instead, the Laplace transform method is used to obtain the transfer function (the output transform divided by the input transform) of the system. The poles and zeros of the transfer function define the system's behavior. Typically, the system is unstable if there are poles in the right half of the s plane (where s is the Laplace transform variable). The use of the computer greatly reduces the tedium of the design process by displaying on a monitor pole-zero plots on the s plane as parameters are varied, the frequency response, the transient response, and other items of interest. The possibility of poles in the right half of the s plane can also be determined by using the Routh-Hurwitz criteria. *See* CONTROL SYSTEM STABILITY; LAPLACE TRANSFORM.

Computer-aided design is particularly helpful in the design process because circuit components such as transistors, resistors, and capacitors can have a significant variation around their nominal values. For example, resistors with a 20% tolerance might be used because they are inexpensive. The amplifier can be described by using the state-variable method in which a system of first-order differential equations for the circuit is determined. The aforementioned properties of stability, response, and so forth, can be determined. *See* LINEAR SYSTEM ANALYSIS.

The computer also provides great help in determining the effects of adding compensating networks to improve the frequency response, the effects of different feedback configurations, and so forth. Thus, the time needed for breadboarding and experimenting is reduced. However, the ultimate test is the performance of the final circuit since the computer prediction is only as good as the equations that model the circuit. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Input and output impedances. The above example (see illus.) shows voltage feedback where a function of the output voltage is fed back to the input. Current feedback, where a voltage proportional to the output current or some function of the output current is fed back, may also be used. Furthermore, an amplifier may have both voltage and current feedback.

The input and output impedances of a feedback amplifier depend upon whether the output current or voltage is sampled and upon whether this signal is fed back in series or in parallel with the input excitation. It is possible for the impedance with feedback to be greater than or less than what it was before the feedback was added. An emitter-follower is an example of a feedback amplifier where the output impedance has been reduced to a very low value and the input impedance has been increased considerably.

Oscillator feedback. An oscillator can be viewed as a feedback amplifier with an input voltage supplied from its output. Referring to the illustration, this would mean that $\mathbf{E}_s$ equals zero and $\mathbf{E}_i$ equals $\mathbf{E}_f$.

From this viewpoint the condition for oscillation at a frequency f_o is (to a first and usually very good approximation) that $\mathbf{A}\beta = 1$ at $f = f_o$. This means that the feedback must be positive.

Servomechanism feedback. The purpose of feedback in a servomechanism is to obtain better control of the system. As an example, consider a position control system which is used to position an object at a point determined by a reference signal.

The command to move the object is derived by comparing the reference signal with a signal indicating the instantaneous location of the object. The command signal is an error signal derived from the comparison of the actual and desired signals; any error signal drives the system in such a direction as to reduce the error, that is, to make the actual position agree with the desired reference position.

If feedback were not used in the position control system, a precisely calibrated control device would be needed to position the object in each position dictated by the reference. In general the required control could not be built with sufficient accuracy. *See* CONTROL SYSTEMS; SERVOMECHANISM.

Harold F. Klock

Bibliography. P. M. Chirlian, *Analysis and Design of Integrated Electronic Circuits*, 2d ed., 1987; J. D'Azzo and C. Houpis, *Linear Control System Analysis and Design*, 3d ed., 1988; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Feeding mechanisms (invertebrate)

The feeding methods of invertebrates are as diverse as the invertebrates themselves, which are adapted to all kinds of habitats, in freshwater, in the sea, and on land. The feeding mechanisms are best classified by the method used: browsing, suspension feeding, deposit feeding, carnivorous, and phytophagous (plant-eating). An alternative classification often adopted, but perhaps less satisfactory, may be based on the size of particles ingested. Thus, the same invertebrate may be described either as microphagous (feeding on minute organisms) or as dependent on substances in solution. Both classification schemes may be further subdivided. Carnivorous feeders, for example, include predators and animal parasites; both share dependence on other (living) animals as a food source. Some methods will be restricted to particular habitats. Suspension feeders, for example, can only be aquatic, while the phytophagous habit may be found wherever edible plants occur. The feeding mechanisms employed will depend, at least partly, on the habit and habitat of the plants utilized.

The insects are by far the most abundant in species and variety. Their feeding mechanisms depend on modification of the mouthparts to suit different foods and on adaptation of the gut to cope with different diets. The mouthparts comprise three pairs of appendages borne on the head—the mandibles, the maxillae, and the labium, each capable of adaptation for biting, piercing, lapping, or sucking. Different adaptations are shown by the mouthparts of, say, a locust for biting off pieces of living plants, a mosquito for piercing mammalian skin and sucking the blood from dermal capillaries, an aphid greenfly for tapping plant sap, and a housefly for lapping fluids, but all may be referred to the same basic plan. Both aphids and locusts are phytophagous in that they depend on plants for food, but their feeding mechanisms are entirely different. *See* INSECTA.

Browsers. The gastropod mollusks afford good examples of browsing invertebrates. The buccal mass houses a rasping radula that acts as a food grater, removing the layer of encrusting organisms (mainly algae) from the rocks. The rasp lines of limpets can often be seen on algae-encrusting rocks in the intertidal zone. The chitons also have a radula and feed by browsing. The radula is a narrow band of tissue bearing rows of minute chitinous teeth; it is produced continuously to replace the anterior part as it wears away. The different dental arrangements are characteristic of each group. The radula is manipulated by a complex system of muscles, for the radula serves not only for rasping the food off the rocks but for conveying it into the buccal cavity. *See* MOLLUSCA.

Carnivores. It is not only the more active invertebrates such as squids and octopus or the planktonic arrowworms (Chaetognatha) that pursue prey; jellyfish and sea anemones catch the animals when they brush against them, and eat them. Nor is it only the larger invertebrates that are predatory. The ciliate

Didinium is voracious in attacking other ciliate protozoa.

The cnidarians are characteristically armed with stinging cells (nematocytes) containing nematocysts (capsules with eversible threads) that may both paralyze and catch fishes and other animals that touch them. The same basic mechanism is adapted by jellyfishes, by the pelagic Portuguese man-of-war (*Physalia*), anemones, and, on a smaller scale, the individual hydranths of colonial hydroids and the solitary freshwater *Hydra*. There is great variety and specialization in the different types of nematocysts for stinging and capture. The comb jellies (Ctenophora) "fish" in the same manner, though their cells (collocytes) are different. After the food is secured (by adhesion), muscular contraction of the tentacles serves to bring the prey to the mouth. *See* CNIDARIA; CTENOPHORA.

Few annelid worms are predators, apart from the leeches, though some polychaetes do prey on other invertebrates, which they catch by means of their jaws and eversible proboscis. Nemertine ribbon worms feed on other animals, mostly when dead or moribund, by protrusion and withdrawal of their long proboscis. Advanced leeches (such as the medicinal leech) pierce the skin of the host by means of three bladelike jaws edged with minute teeth, sucking the blood which is prevented from clotting by anticoagulants in the saliva. Such predators (or parasites) take large meals at infrequent intervals and have large crops for storage. Some leeches such as *Trochaeta* ingest whole earthworms, and some freshwater leeches ingest other small invertebrates. Mollusks such as *Urosalpinx* prey on barnacles, gastropods, or bivalves, boring through the shells by a combination of chemical and mechanical means. The process is often lengthy and laborious but is rewarded by a substantial supply of food. Preliminary digestion may be external, the partially digested tissue being sucked up by means of a tubular proboscis. *See* ANNELIDA; MOLLUSCA.

Some sea stars use their tube-feet to pull apart the valves of mussels, the star everting part of its stomach between the valves to digest and absorb the meat. *See* ECHINODERMATA.

Scavengers. Many invertebrates are scavengers, depending on animal or plant remains, or on animals in a moribund condition. Many crustaceans are scavengers, including isopods and amphipods; and many of the larger crabs also fall into this category. The multiple mouthparts of such animals are adapted to carry out different parts of the feeding process. Large pieces of food may be grasped by pincerlike (chelate) appendages and passed via other pairs of appendages that comminute the food into fragments of a size suitable for ingestion. *See* CRUSTACEA.

Suspension feeders. Feeding on particles in suspension is particularly widespread among intertidal invertebrates, for the action of the waves and the high productivity of coastal waters ensure a relatively abundant supply of detritus and living organisms. Some filter feeders have remarkably efficient filtering mechanisms (**Fig. 1**), and the adoption of suspension feeding is often combined with a sedentary life. The mechanisms employed may be subdivided into setose, ciliary, and mucous-bag feeding.

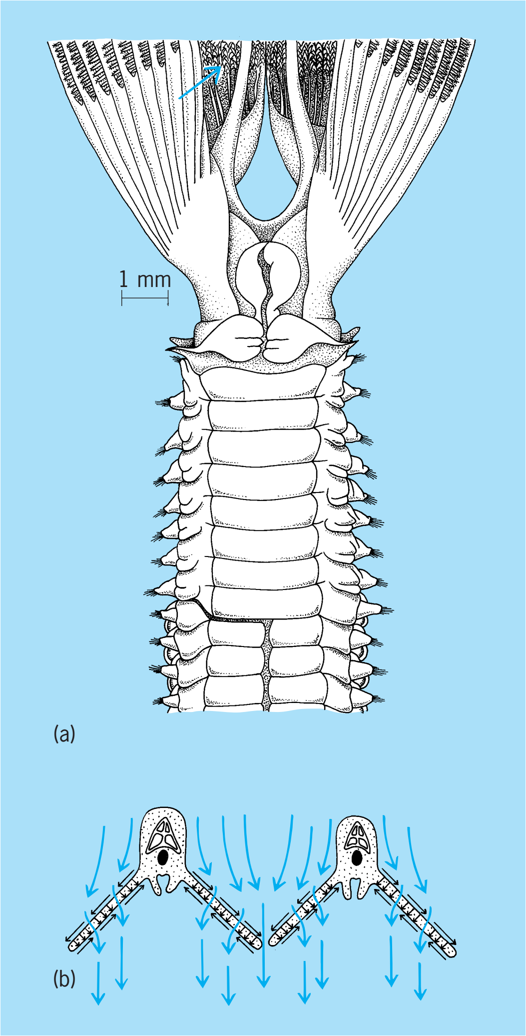


Fig. 1. ***Sabella pavonina*. (*a*) Ventral view of base of tentacular crown and first 12 body segments. (*b*) Diagrammatic transverse section of two tentacles, as indicated by arrow on *a*, to show the direction of the flow of water entering the crown (colored arrows) and the direction of the ciliary beat on the pinnules (black arrows). The longitudinal food groove of each tentacle is overarched by expanded bases of pinnules; particles from pinnules enter it between these bases. (*After V. Fretter and A. Graham, A Functional Anatomy of Invertebrates, Academic Press, 1976*)**

Setose appendages. Suspension feeding by means of setose appendages is characteristic of many crustaceans such as acorn barnacles, which secrete a limy carapace and use six pairs of appendages (the cirri) to strain particles from the water and direct them to the mouth. In strong currents the cirri need only be held out into the current, the particles are caught on the setae, and the cirri are contracted from time to time to convey the catch to the mouth. Some crabs, such as the porcelain crabs, also depend on particles in suspension. Various appendages participate in the filtering process, but in general setose extensions of mouthparts or maxillipeds are used. Some sand crabs and amphipods have some appendages adapted for filtering and others adapted for dealing with larger particles so that they can utilize a wide range of food as opportunity permits.

Ciliary mechanisms. The fan worms such as *Sabella* depend on currents drawn through a crown of pinnate filaments by long cilia along the edges of the filaments, particles being trapped by mucus produced by glandular cells and conveyed toward the mouth along defined tracts by motion of shorter cilia. The cilia beat in particular directions so that a stream of trapped particles converge from each pinnule to the base of the filament, where they are sorted according to size. Unsuitable particles are rejected by other ciliary currents sweeping them away from the crown (Fig. 1).

Some intertidal gastropods have adapted their respiratory currents into feeding currents and their gills into ciliary filters, a mechanism that may have evolved from the necessity not only of keeping the gills irrigated but of keeping them clean. Gastropods such as the slipper limpet, *Crepidula fornicata*, have therefore become sedentary, for by transferring the gill cleanings to the mouth there is no necessity for locomotion. The Chinaman's-hat limpet (*Calyptraea*) feeds in a similar manner. So does the tall *Turritella* to a certain extent, but since much of its food is sucked up from the surface of the substratum, it is principally a deposit feeder.

Clams and oysters are the most efficient filterers in that the water is passed through the gills, the particles being sorted mainly by size on special flap-shaped palps by the mouth. The enormous mussel (*Mytilus*) banks on rocky coasts bear testimony to the amount of food in suspension in such places. Mussels avoid being swept away by secreting threads (byssus), but rock oysters and many other filterers cement themselves to rocks and pilings in order to take advantage of these nutritive currents.

The protochordate sea squirts have achieved the same type of feeding by other means. The filter is a branchial basketlike elaboration of pharyngeal gill slits. The feeding current is taken in through the mouth and passed through the slits and out through the atrium and exhalant aperture (**Fig. 2**). The food is trapped in a mucous string produced from a ventral gutter (endostyle), the mucus flowing across the basket to the esophagus. Like many other filterers, the sea squirts have become sedentary.

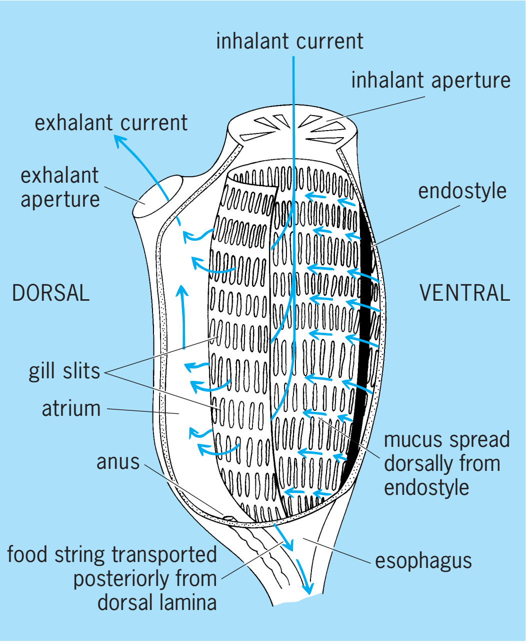


Fig. 2. Diagrammatic sagittal section of *Clavelina* showing the branchial basket and feeding currents. The side of the tunic and part of the branchial basket are cut away. (*After R. C. Newell, Biology of Intertidal Animals, Marine Ecological Surveys Ltd., Faversham, Kent, England, 1979*)

The bryozoans, common below tide on many shores, have developed a filtering crown of tentacles, the lophophore, superficially like that of hydroid cnidarians; and they are also sedentary and colonial. The bryozoan inhalant current is directed by cilia down into the center of the crown and out between the bases of the tentacles, the food particles being directed along prescribed paths to the mouth. The solitary phoronids are similar; and the brachiopod lamp shells, so numerous in earlier times, have a lophophore shaped like a double horseshoe of numerous slender filaments within the lower valve of the Roman lamp–shaped shell. The mechanism is functionally like that of a mussel, but the construction is entirely different. *See* BRACHIOPODA; BRYOZOA; PHORONIDA.

Mucous bag. Secretion of mucus, denied to the crustaceans but an indulgence in soft-bodied invertebrates, is a common means of trapping food. In a few animals the mechanism is more complex. In general, nets or bags of mucus are spun from the mouth; when the device is full, both the net and the catch are eaten. The net is filled by a water current that may be actively provided by muscular or other means, or the feeder may depend on the natural current. Some insect larvae use this method in freshwaters, while the annelid *Chaetopterus* and the echiuroid *Urechis* do so in the sea. Mucous-bag filters trap minute particles. *Chaetopterus* (**Fig. 3**) secretes a parchmentlike tube whereas *Urechis* lies in a burrow, but both make a current through a secreted mucous bag. *Chaetopterus* produces its current with muscular balers or "fans," *Urechis* by peristalsis of the body wall. In *Chaetopterus* the bag is subtended by two armlike extensions that secrete it, the bag rolled up in a cup-shaped structure on the segment behind, and the bolus passed forward along a groove to the mouth.

Integumentary absorption. Some invertebrates absorb all their nutrients through the integument. This is true not only of many parasites bathed in the nutritious media of their hosts' body fluids or gut contents but also of the pogonophoran annelids, which live in cold deep-water marine deposits and have neither mouth nor gut. Slow and lethargic, they live in permeable horny tubes through which their nutrients are actively absorbed from solution. Some other invertebrates are capable of utilizing

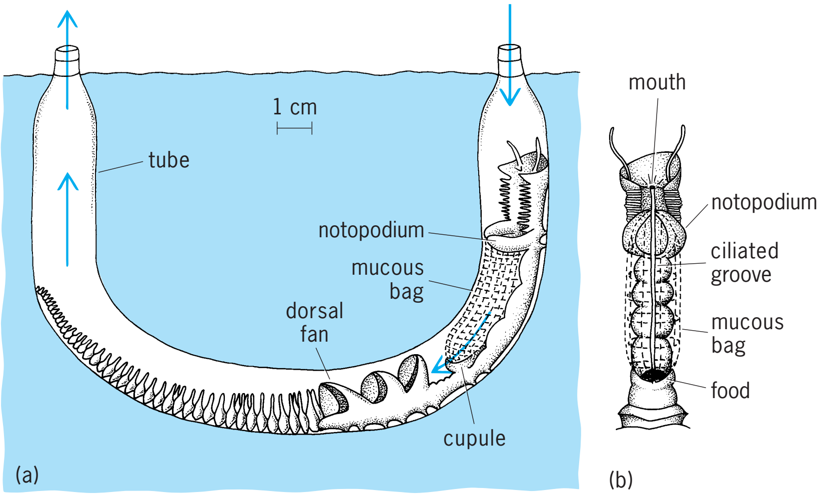


Fig. 3. ***Chaetopterus*. (*a*) Feeding with a mucous bag, the arrows indicating water current. (*b*) Dorsal view of the anterior end. (*After V. Fretter and A. Graham, A Functional Anatomy of Invertebrates, Academic Press, 1976*)**

dissolved substances, amino acids or complex carbohydrates, by active uptake. This form of nutrition has often been overlooked, but it may form a small but significant proportion of the energy budget. *See* POGONOPHORA.

Deposit feeders. Many invertebrates feed on deposits of mud and sand. The nutritional content may be relatively low, but it is usually highest at the surface where there is a microflora of photosynthetic organisms and where detritus is also deposited. There are various mechanisms for collecting the most nutritious surface layer, and guts of such feeders are specialized for coping with large volumes of deposit continuously delivered.

In the intertidal zone the polychaete annelids afford a variety of examples of deposit-feeding mechanisms. The lugworm (*Arenicola*) lies at a depth of 8–16 in. (20–40 cm) in an L-shaped burrow, with the head in the toe of the L, and feeds on sand. As it eats, sand caves in from above, forming hollows on the surface. From time to time, the worm backs up the vertical shaft to form the castings of sand that has passed through the gut. The cycles of eating and defecation continue with clockwork regularity, for there is little to disturb the even tenor of its ways. Other polychaetes living beneath the surface, or in crevices for protection, have developed tentacles that can be extended onto the surface. *Terebella* is one such, with numerous extensible tentacles arising, medusalike from the head, each provided with muscles and a ciliated groove to assist in conveying morsels to the mouth. Other invertebrates, such as ampharetid worms and sea cucumbers (*Holothuria*), have shorter sticky tentacles that are licked clean of adherent particles by the mouth. *See* POLYCHAETA.

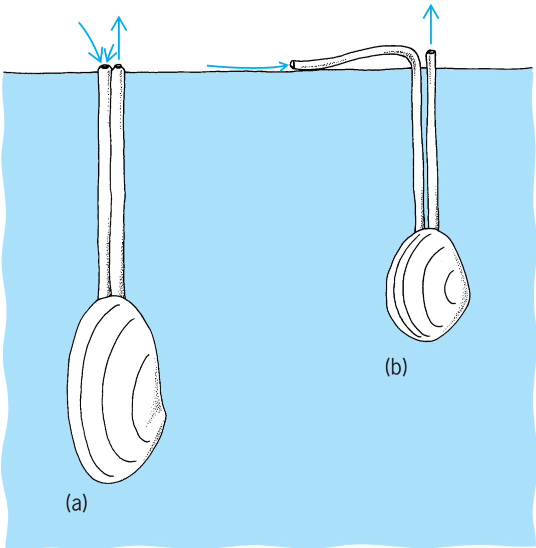


Fig. 4. Use of siphons in (*a*) a suspension-feeding bivalve (*Mya arenaria*) and (*b*) a deposit-feeding bivalve (*Scrobicularia plana*). Arrows indicate the inhalant and exhalant streams. (*After R. C. Newell, Biology of Intertidal Animals, Marine Ecological Surveys Ltd., Faversham, Kent, England, 1979*)

Most bivalves which are not filterers ingest the surface deposits by means of the inhalant siphon, which acts rather like a domestic vacuum cleaner. The suspension-feeding *Mya* lies well buried in the mud, with inhalant and exhalant siphons externally joined together and extending to the surface (**Fig. 4*a***). The deposit-feeding *Scrobicularia* has the inhalant and exhalant siphons separated, enabling the inhalant tube to search a wide area of surface while the rest of the animal remains at a safer depth (Fig. 4*b*). Many gastropods, like the tiny *Hydrobia*, feed on surface mud. The larger pelican's-foot snail (*Aporrhais*) has an elaborate mechanism for rasping off organisms from the deposits which are taken in from the surface by means of an inhalant current. Both inhalant and exhalant currents are produced by the proboscis and consolidated by the mucus it secretes. As in *Arenicola*, the surface is eaten while the animal remains concealed. *Corophium*, an amphipod that burrows in fine marine mud, takes the rich surface into the opening of the burrow by means of its long second antennae. The beating of the more posterior appendages (pleopods) creates a current that carries the raked-up material into a setose sorting area near the mouth.

In terrestrial environments there are, of course, no filter feeders, and deposit feeders live mainly in the soil to avoid light and water loss as well as predators. Terrestrial vegetation provides food for browsers such as slugs and snails and for numerous insect larvae with mouthparts adapted to deal with particular plants. Plant litter in various stages of decay provides food for other invertebrates both on and beneath the surface. Such invertebrates have not only mouths and teeth to deal with the food but also gut floras to assist digestion. Cellulose and wood are often broken down with the aid of gut flagellates. *See* DIGESTION (INVERTEBRATE). R. Phillips Dales

Bibliography. R. S. K. Barnes, P. Calow, and P. J. W. Olive, *The Invertebrates: A Synthesis*, 3d ed., 2001; R. C. Brusca and G. J. Brusca, *The Invertebrates*, 2d ed., 2003; P. A. Meglitsch, *Invertebrate Zoology*, 3d ed., 1991; V. Pearse et al., *Living Invertebrates*, 1989; C. L. Prosser, *Comparative Animal Physiology*, 4th ed., 1991; C. M. Yonge, Feeding mechanisms in the invertebrates, *Biol. Rev.*, 3:21–70, 1928.

Feldspar

Any of a group of aluminosilicate minerals whose crystal structures are composed of corner-sharing [AlO_4] and [SiO_4] tetrahedra linked in an infinite three-dimensional array, with charge-balancing cations [primarily sodium (Na), potassium (K), and calcium (Ca)] occupying large, irregular cavities in the framework of the tetrahedra. Collectively, the feldspars constitute about 60% of the outer 8–10 mi (13–17 km) of the Earth's crust. They are nearly ubiquitous igneous and metamorphic rocks, and are a primary constituent of arkosic sediments derived from them. The importance of the many feldspars that occur so widely in igneous, metamorphic, and some sedimentary rocks cannot be underestimated, especially from the viewpoint of a petrologist attempting to unravel earth history. *See* ARKOSE; MINERALOGY; PETROLOGY; SILICATE MINERALS.

With weathering, feldspars form commercially important clay materials. Economically, feldspars are valued as raw material for the ceramic and glass industries, as fluxes in iron smelting, and as constituents of scouring powders. Occasionally their luster or colors qualify them as semiprecious gemstones. Some decorative building and monument stones are predominantly composed of weather-resistant feldspars. *See* CLAY MINERALS; IGNEOUS ROCKS; METAMORPHIC ROCKS.

Chemical Composition

The general formula AT_4O_8 characterizes the chemistry of feldspars, where T (for tetrahedrally coordinated atom) represents aluminum (Al) or silicon (Si). The A atom is Ca^{2+} or barium (Ba^{2+}) for the $[Al_2Si_2O_8]^{2-}$ alkaline-earth feldspars and Na^+ or K^+ for the $[AlSi_3O_8]^-$ alkali feldspar series of solid solutions and mixed crystals. A complete range of compositions is observed in the plagioclase feldspar series, $Na_yCa_{1-y}Al_{2-y}Si_{2+y}O_8$ ($0 \leq y \leq 1$), and in a somewhat analogous, though rare, $K_xBa_{1-x}Al_{2-x}Si_{2+x}O_8$ series called hyalophanes. The $BaAl_2Si_2O_8$ end member is called celsian (symbolized Cn).

The majority of natural feldspars have the approximate formula

$$K_xNa_yCa_{1-(x+y)}Al_{2-(x+y)}Si_{2+(x+y)}O_8$$

where $0 \leq (x+y) \leq 1$, and x is equal to the mole fraction of the potassium feldspar end member (Or, for orthoclase), y is equal to the mole fraction of the sodium feldspar (Ab, for albite), and $1-(x+y)$ is equal to the mole fraction of the calcium feldspar (An, for anorthite).

Knowledge of a feldspar's composition (**Fig. 1**) and its crystal structure is indispensable to an understanding of its properties. However, it is the distribution of the Al and Si atoms among the available tetrahedral sites in each chemical species that is essential to a complete classification scheme, and is of great importance in unraveling clues to the crystallization and thermal history of many igneous and metamorphic rocks.

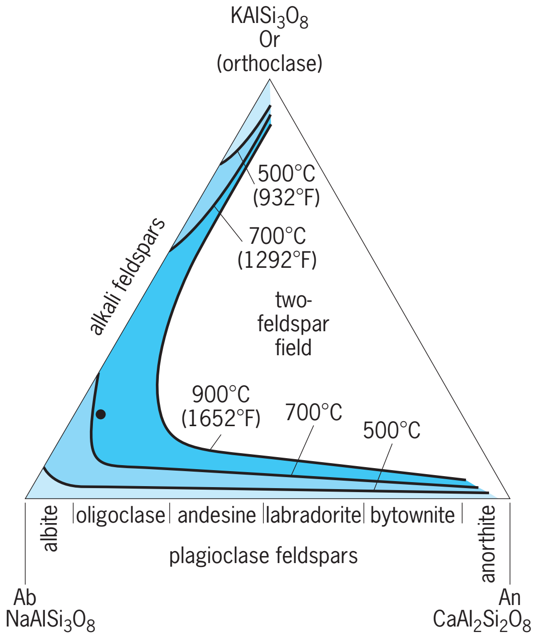


Fig. 1. The feldspar Or-Ab-An ternary as determined at a water pressure of 1000 atm (6.9 MPa). The solid circle represents a feldspar quenched above 700°C (1290°F) whose composition is $Or_{20}Ab_{75}An_5$ or $K_{0.20}Na_{0.75}Ca_{0.05}Al_{1.05}Si_{2.95}O_8$. The range of feldspar compositions is shown as a function of increasing temperature in the successively less shaded areas.

Crystal Structures

The structures of all feldspars share similar characteristics. The primary structural units are four-membered rings of AlO_4 and SiO_4 tetrahedra linked to similar rings by sharing certain oxygen atoms (**Fig. 2**). In an idealized concept (Fig. 2*b*), crankshaft-like chains are shown as extending through the crystal parallel to the *a* unit cell dimension, and they occur in pairs that are mirror images (or approximate mirror images) of one another, across the (010) crystallographic plane. Parallel to the *b* unit cell dimension, the chains are joined together by the T-O-T, bonds, and in the *c* direction slabs of chain pairs are joined by T-O-T bonds through oxygen.

Large A cations are located between the (001) slabs; Na^+ and Ca^{2+} are bonded to six or seven oxygen atoms and K^+ to nine oxygen atoms. The prominent (010) and (001) cleavages of feldspars are expressions of the weakness of the A-O bonds.

Complete disorder of Al and Si is virtually unknown in natural potassium feldspars. At lower temperatures, Al increasingly concentrates in T_1 sites and Si in T_2 sites, but this ordering process is limited by the number of Al atoms per formula unit. If all the Al is to end up in a single T_1 site (which in nature is the

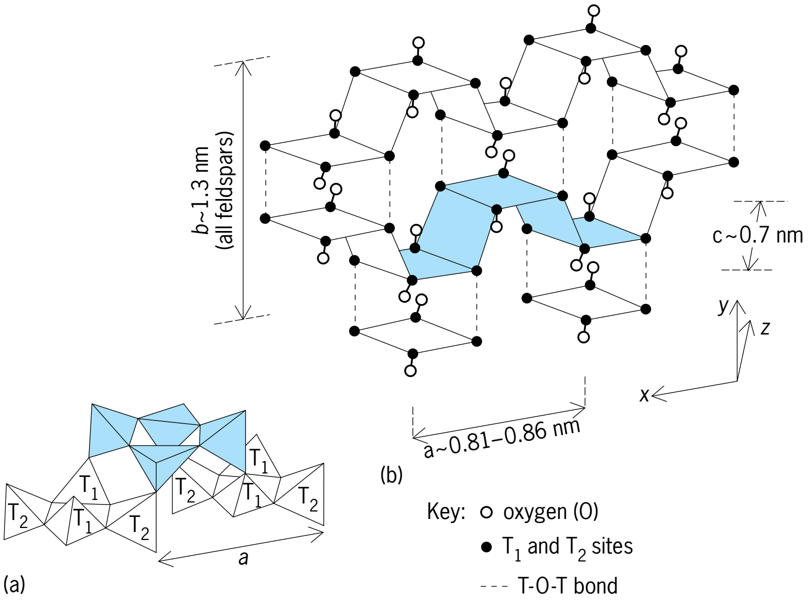


Fig. 2. Feldspar structure. (*a*) Perspective view showing the linkage of AlO_4, and SiO_4 tetrahedral units, T_1 and T_2, into fourfold rings; oxygen atoms are at the corner of the tetrahedra. The shaded tetrahedral ring is related to the unshaded ones by an *a*-glide plane parallel to (010). (*b*) Expanded and idealized representation, showing both T_1 and T_2 sites as black dots. These are a pair of mirror- or pseudomirror-related crankshaftlike chains. The shaded area is that portion of this drawing that is projected onto the (010) plane in Fig. 3*a*.

configuration with the lowest free energy for alkali feldspars), the T_1 sites must become symmetrically nonequivalent (**Fig. 3***b*). An O suffix is arbitrarily assigned to the T_1 site into which Al is concentrated; the other site is called T_1m because in the feldspar structure it is related to T_1O by a pseudomirror (m) operation. The T_2 sites are likewise differentiated. There is distortion of the fourfold ring, and the monoclinic space group (C2/m) feldspar becomes triclinic ($C\bar{1}$).

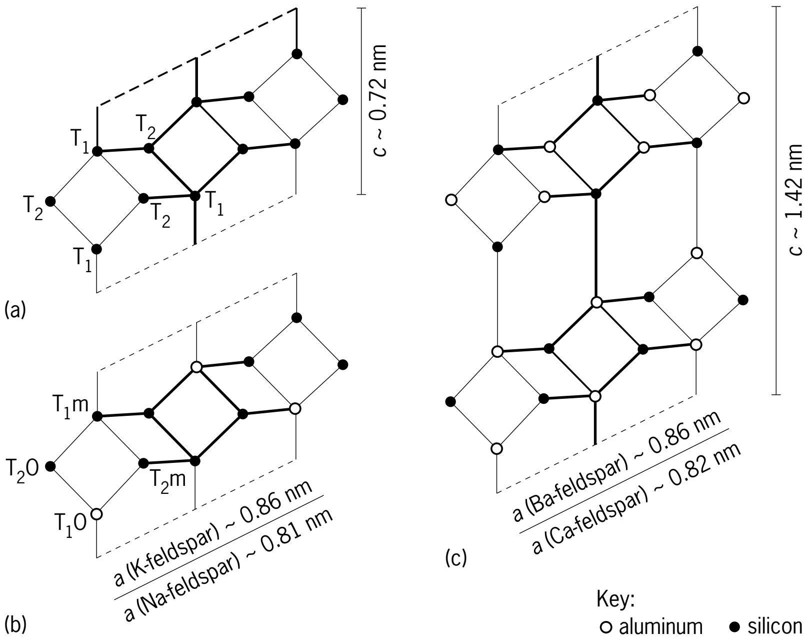


Fig. 3. Feldspar structures showing the shaded portion of Fig. 2*b* projected onto (010). (*a*) Monoclinic sanidine and monalbite. (*b*) Triclinic low albite and low microcline [Al ordered into the T_1O sites]. (*c*) Monoclinic celsian or triclinic anorthite [note the doubled *c* cell dimension]. The angle between the *a* and *c* unit cell edges is called β; $\beta \sim 116^\circ$ in all feldspars.

Feldspars containing more than a few percent of divalent cations may have very complex crystal structures or exsolve into two phases, or both, primarily because Al:Si ratios depart from the 1:3 ratio of the relatively simple alkali feldspars discussed above. In nature, the Al and Si atoms of the celsian and anorthite end members alternate regularly in the tetrahedral framework. Such an ordered Al,Si distribution (Fig. 3*c*) requires that the *c* cell dimension be doubled to about 1.4 nm from the approximately 0.7-nm characteristic of the C2/m and $C\bar{1}$ feldspars (Fig. 3*a*,*b*). Celsian is monoclinic (space group, I2/c). Anorthite, whose A site is occupied by a much smaller atom than that in celsian [radius of Ca $\simeq$ 0.10 nm, that of Ba $\simeq$ 0.14 nm], has a collapsed triclinic framework (space group $I\bar{1}$ or $P\bar{1}$). *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Nomenclature

The feldspar minerals are broadly divided into two classes, alkali feldspars and plagioclase feldspars (**Tables 1** and **2**).

Alkali feldspars. Names are assigned to the polymorphs of $KAlSi_3O_8$ and $NaAlSi_3O_8$ in accordance with their symmetry and the Al content of their tetrahedral sites. *See* ALBITE; ORTHOCLASE; MICROCLINE.

The names pericline and cleavelandite apply to particular morphological varieties of albite, and adularia is a variety of K-rich feldspar. Anorthoclase is

TABLE 1. Alkali feldspars, in order from high- to low-temperature forms

Potassium-rich feldspars	
Monoclinic (space group C2/m)	
High sanidine	$0.5 \leq 2t_1 \leq 0.67$
Low sanidine	$0.67 < 2t_1 < 0.75$
Orthoclase	$0.75 \leq 2t_1 \leq 1.00$
Triclinic (space group C1)	
Intermediate microcline	$1.0 < t_1o > t_1m > t_2o = t_2m$
Low microcline or maximum microline	$t_1 \simeq 1.0; t_1m = t_2o = t_2m \simeq 0.0$ (fully ordered)
Sodium-rich feldspars	
Monoclinic (space group C2/m)	
Monalbite	$0.5 \leq 2t_1 \leq \sim 0.6$ (exists only at temperatures above 980°C or 1800°F)
Triclinic (space group C1)	
Analbite	$0.5 \leq 2t_1 \leq \sim 0.6$ (the lower-temperature form of monalbite)
High albite	$t_1o > t_1m \geq t_2o = t_2m$ (highly disordered; all *t* values near 0.25)
Intermediate albite	$t_1o > t_1m \geq t_20 = t_2m$ (probably not stable in nature)
Low albite	$t_1o \simeq 1.0; t_1m = t_2o = t_2m \simeq 0.0$ (fully ordered)

TABLE 2. Plagioclase feldspars

Name	Composition
Sodic (acid) plagioclases	
Albite	An_0–An_{10}
Oligoclase	An_{10}–An_{30}
Andesine	An_{30}–An_{50}
Calcic (basic) plagioclases	
Labradorite	An_{50}–An_{70}
Bytownite	An_{70} An_{90}
Anorthite	An_{90}–An_{100}

a triclinic solid solution of composition $Or_{37}Ab_{63}$–Or_0Ab_{100} containing up to 10 mol % anorthite, or more. *See* ANORTHOCLASE.

Ternary feldspars are those containing at least 10% of all three end members; in nature most anorthoclases and ternary feldspars will have exsolved into two discrete phases.

Plagioclase feldspars. Nomenclature of the plagioclases is based on their compositions (Fig. 1). Plagioclases containing significant amounts of exsolved K-rich feldspar are called antiperthites. It is only in once-molten rocks quenched at very high temperatures that the full range of so-called high plagioclases exist as simple solid solutions. With very slow cooling over millions of years, complex textures develop in most feldspar crystals as a coupled NaSi, CaAl ordering. *See* ANDESINE; BYTOWNITE; LABRADORITE; OLIGOCLASE; SOLID SOLUTION.

Properties

The variable properties of feldspars are determined by their structure, symmetry, chemical composition, and crystallization and subsequent history of phase transformation, exsolution, and alternation or deformation.

Color. Very few feldspars are transparent and colorless; many are white or milky due to internal reflections of light from inclusions, exsolution interfaces, and fracture or cleavage surfaces. In fact, the name for the sodium feldspar is albite (from the Latin *albus*, white). Fine-scale alteration to sericite or clay produces a cloudy appearance, and if tiny dispersed crystals of hermatite (Fe_2O_3) are present, a cloudy pink to orange-brown to brick-red color results. It is thought that black-clouded feldspars (most often calcic plagioclases) are caused by minute, dispersed iron oxides or iron-titanium oxides. Yellow feldspars contain minor Fe^{3+}. Blue to green feldspars (including amazonite) contain Pb^{+} and structurally combined H_2O. Smoky feldspar, like smoky quartz, results from ionizing radiation.

Oriented intergrowths resulting from exsolution produce a variety of effects due to scattering, reflection, and interference of incident light. Aventurine and sunstone are feldspars in which thin flakes of hematite have exsolved in a particular crystallographic orientation and reflect light as glittering rosy or gold schiller (play of color); copper platelets produce pink schiller. Sometimes, oriented mica flakes or metal oxide platelets and needles cause a silvery sheen or chatoyancy. Exsolution textures in cryptoperthitic alkali feldspars may result in an ethereal blue to whitish appearance. These feldspars are known as moonstones. In the three regions of exsolution in the plagioclase subsolidus, the resulting two-phase intergrowths are essentially planar. Light interference, like that from an oil slick on water, produces brilliant red, yellow, green, and blue colors.

Mechanical properties. Plagioclases are slightly harder (6–6.5) on Mohs scale than K-rich feldspars (6). Feldspars are brittle and, when broken, cleave along the (001) and (010) crystallographic planes. Both cleavages are parallel to the crankshaftlike chains (Fig. 2), and both represent planes of weak A–O bonding. Quartzlike conchoidal fracture is occasionally seen in crystals that have a glassy appearance.

Morphology. Feldspars that have grown relatively unimpeded in fluids or cavities usually have the following primary crystal faces: $\{110\}$, $\{1\bar{1}0\}$, $\{010\}$, $\{001\}$, $\{20\bar{1}\}$, and $\{10\bar{1}\}$; faces of lesser importance are $\{130\}$, $\{1\bar{3}0\}$, $\{1\bar{1}\bar{1}\}$, and $\{1\overline{11}\}$ (**Fig. 4**). The principles controlling crystal growth and favoring the development of one or more crystal faces over others are complex.

Twinning. As a group, no minerals have more complex and ubiquitous twins than feldspars. If only two (or a few) individuals are observed, the twins are called simple twins (**Fig. 5*a*, *b***). Complex twins result when two twin laws with the same composition plane are simultaneously operative in a single crystal; Albite-Carlsbad is a relatively common example in plagioclases. Albite and Pericline twins are nearly always multiple or polysynthetic twins (Fig. 5*c*, *d*), and they are observed only in triclinic feldspar. The (010) composition planes of albite twins often show up as straight lines (actually shallow grooves) on (001) faces (Fig. 5*c*). Traces of Pericline twins are seen on (010) cleavage planes as parallel straight lines (Fig. 5*d*), inclined to the trace of the (001) cleavage.

In practice, all types of twins are most easily observed and identified with a petrographic microscope, which is widely used in the study of rocks and minerals. Genetically there are three main categories of twins: (1) growth twins, (2) inversion or transformation twins, and (3) deformation or mechanical twins. *See* PETROGRAPHY.

Density. Densities are measured in terms of mass per unit volume. Densities are near 2.56 g/cm^3 for

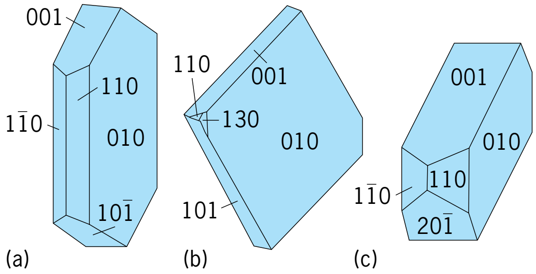


Fig. 4. Various feldspar morphologies: (*a*) blocky, typical of microcline; (*b*) platy, typical of quenched feldspars called microlites; (*c*) prismatic, typical of orthoclase.

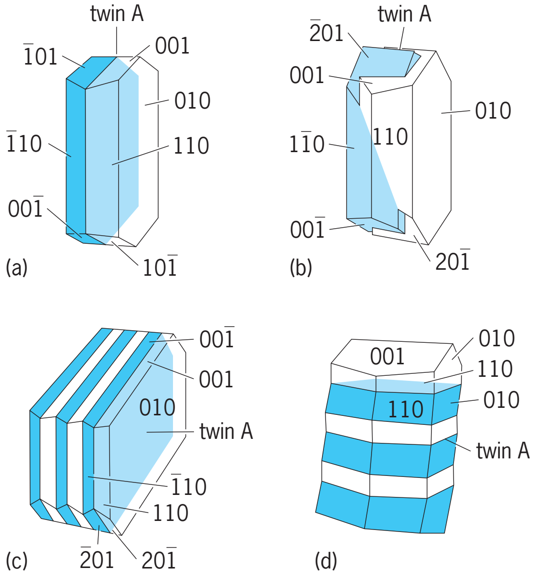


Fig. 5. Simple and multiple twins in feldspars. (*a*) Carlsbad (contact), (*b*) Carlsbad (interpenetrant), (*c*) albite, (*d*) pericline. (*After T. Zoltal and J. H. Stout, Mineralogy: Concepts and Principles, Burgess, 1984*)

K-feldspars and vary regularly with composition between 2.62 g/cm^3 for Na-feldspar and 2.78 g/cm^3 for Ca-feldspar. Ba-feldspar has a density of 3.37 g/cm^3.

Optical properties. Feldspars are optically biaxial, exhibiting three refractive indices whose values and vibration directions within the crystal can be diagnostic of composition and degree of Al,Si order-disorder.

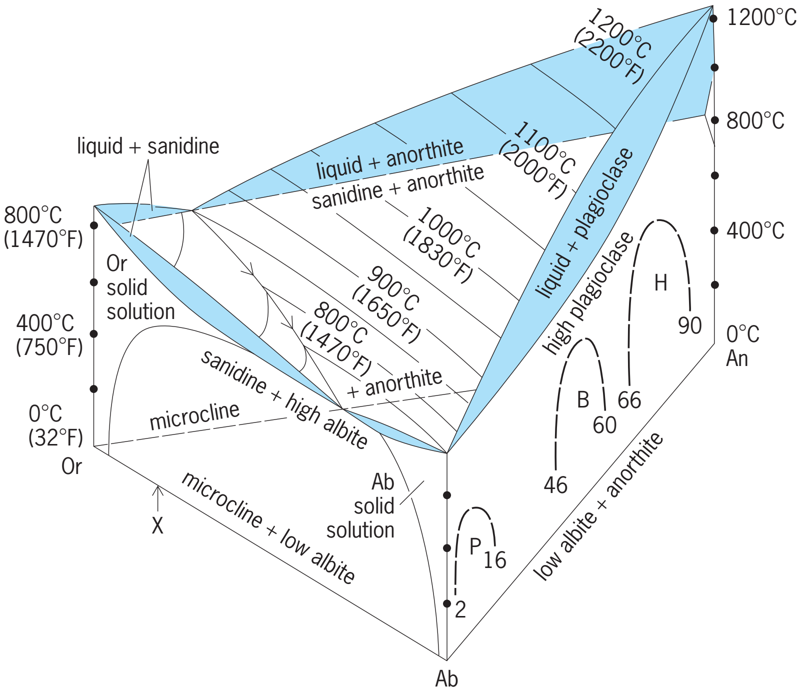


Fig. 6. Temperature-composition diagram for the Or-Ab-An system at a water pressure near 5000 kg · cm^{-2} (490 MPa). Details of the interior of this pseudo phase diagram are complex and have been omitted. P = peristerite intergrowth. B = Bøggild intergrowth. H = Huttenlocher intergrowth. Numbers at intergrowths represent approximate ranges (in mol % An) over which exsolution occurs. Temperatures are given in degrees Celsius (Fahrenheit).

Although these optical properties are relatively complicated, they are widely and routinely used, together with extinction angles of twins, by geologists in their study of the history of crystallization, metamorphism, or mechanical deformation of rocks in which these ubiquitous feldspars occur.

Melting properties. In a dry (water-free) system K-rich feldspar melts incongruently, forming a mixture of leucite ($KAlSi_2O_6$) and silica-rich liquid above about 1075°C (1967°F); albite and anorthite melt congruently at 1118°C (2044°F) and about 1545°C (2813°F), respectively, forming glasses when rapidly quenched. When water is present, the melting temperatures of feldspars are lowered substantially.

Phase Equilibria of the Feldspar Binary Systems

Several isotherms on the Or-Ab-An composition triangle at about 1000 atm (6.9 MPa; Fig. 1) have been determined experimentally. At or above 500°C (930°F), feldspar compositions exist as single phases, that is, as crystalline solid solutions. At or above 700°C (1300°F), all compositions from the 700°C (1300°F) isotherm out to the limiting triangle are single phases, and likewise for the 900°C (1650°F) isotherm. At 900°C (1650°F) or below, bulk compositions in the two-feldspar field are immiscible and form two chemically distinct phases. Phase relations of the K-Na-Ca feldspars are better understood with reference to a diagram (**Fig. 6**) of the three binary systems, Or-An, Or-Ab (alkali feldspars), and Ab-An (plagioclases), versus temperature for an approximate water pressure of 5000 kg cm^{-2} (490 MPa). In such a diagram the top surface (the liquidus) is contoured for temperatures at which a particular composition becomes completely molten. Individual binary diagrams give information on the range of temperatures and compositions at which a crystalline phase is in equilibrium with a liquid. At temperatures below the solidus, only crystalline (solid) phases exist; depending on temperature and composition, there may be a single homogeneous phase or two crystalline solid-solution phases in equilibrium. *See* PHASE EQUILIBRIUM.

System $KAlSi_3O_8$-$CaAl_2Si_2O_8$. A phase diagram, known as a binary eutectic, has been determined experimentally in the laboratory, but it is of little practical importance in nature because bulk compositions in this range do not occur in igneous rocks.

System $KAlSi_3O_8$-$NaAlSi_3O_8$. This system is of considerable importance geologically, and alkali feldspars have been studied in great detail. It has been impossible to duplicate experimentally the textures and structures of many of these feldspars which, in nature, may have had many millions of years to attain their present state. Because of the extremely sluggish kinetics of Al,Si migration in the solid state, true equilibrium assemblages of feldspars in rocks are by no means the rule. However, using combinations of laboratory synthesis and detailed analysis of natural feldspars, petrologists have reached a fairly sophisticated level of understanding of the K-Na system (Fig. 6).

In the pure Or-Ab binary, the structural state and composition of the phase or phases observed will depend on three parameters: (1) the initial bulk composition, (2) the temperature at which the material is observed (or from which it is rapidly quenched to room temperature to facilitate analytical studies), and (3) the annealing conditions, including in particular the rate at which the material has been cooled. Below the solvus the initially homogeneous single crystal begins to exsolve, separating in the solid state into two feldspars. The resulting composite feldspar is called perthite. *See* PERTHITE.

Both the Or and Ab solid-solution phases change not only in composition but also in structural state with slow cooling. Increasingly ordered Al,Si distributions are to be expected (Table 1).

System $NaAlSi_3O_8$-$CaAl_2Si_2O_8$. Crystallization from a liquid in the pure binary system is relatively simple, but in nature, spurts of cooling and reheating, as well as local fluctuations of magma composition (including H_2O and other volatiles), will more than likely produce zoned plagioclase grains. Petrologists often analyze the compositional variation of such feldspars—together with other mineralogical and textural variables—in attempts to decipher the crystallization histories of the rocks in which they occur. *See* MAGMA.

Once all the fluid phases in a rock system are exhausted, zoning patterns may persist throughout geologic time, as long as the rock is not thermally metamorphosed. But whether chemically zoned or homogeneous, the relatively disordered high-plagioclase crystals that form just below the solidus are subject to further structural and textural modification in the solid state.

An Ab-An phase diagram (Fig. 6) can be simplified and yet is adequate to serve as a guide to a discussion of the nature of most plagioclases that are found in igneous and metamorphic rocks.

In nature, most plagioclases in the range An_2 to An_{16} have exsolved into two phases: one is nearly pure low albite, the other is an oligoclase of composition $Ab_{80\pm5}An_{20\mp5}$. The resultant lamellar intergrowths are called peristerites whether or not they display a white, pale blue, or occasionally brilliant range of spectral colors. The colors are due to interference of light reflected from multiple thin films of albite and oligoclase.

Paul H. Ribbe

Bibliography. W. L. Brown (ed.), *Feldspars and Feldspathoids: Structures, Properties and Occurrences*, 1984; I. Parsons (ed.), *Feldspars and Their Reactions*, 1994; P. H. Ribbe (ed.), *Feldspar Mineralogy*, 2d ed., vol. 2 of *Reviews in Mineralogy*, 1993; J. V. Smith, *Feldspar Minerals*, vols. 1 and 2, 1974, rev. 1986.

Feldspathoid

A member of the feldspathoid group of minerals. Members of this group are characterized by the following related features: (1) All are aluminosilicates with one or more of the large alkali ions (for example, sodium, potassium) or alkaline-earth ions (for example, calcium, barium). (2) The proportion of aluminum relative to silicon, both of which are tetrahedrally coordinated by oxygen, is high. (3) Although the crystal structures of many members are different, they are all classed as tektosilicates; that is, they consist of a three-dimensional framework of aluminum and silicon tetrahedrons, each unit of which shares all four vertices with other tetrahedrons, giving rise to a ratio of oxygen to aluminum plus silicon of 2:1. Ions of other elements occupy interframework sites. (4) They occur principally in igneous rocks, but only in silica-poor rocks, and do not coexist with quartz (SiO_2). Feldspathoids react with silica to yield feldspars, which also are alkali–alkaline-earth aluminosilicates. Feldspathoids commonly occur with feldspars.

The principal species of this group are the following:

Nepheline	$KNa_3[AlSiO_4]_4$
Leucite	$K[AlSi_2O_6]$
Cancrinite	$Na_6Ca[CO_3\|(AlSiO_4)_6] \cdot 2H_2O$
Sodalite	$Na_8[Cl_2\|(AlSiO_4)_6]$
Nosean	$Na_4[SO_4\|(AlSiO_4)_6]$
Haüyne	$(Na,Ca)_{8-4}[(SO_4)_{2-1}\|(AlSiO_{46}]$
Lazurite	$(Na,Ca)_8[(SO_4,S,Cl)_2\|(AlSiO_4)_6]$

The last four species (sodalite group) are isostructural, and extensive solid solution occurs between end members; but members of the sodalite group, cancrinite, leucite, and nepheline have different crystal structures.

Nepheline is an essential constitutent of intrusive nepheline syenites and nepheline syenite pegmatites (as in the extensive deposits of the Kola Peninsula, Russia) and of their extrusive equivalents, phonolites. It also occurs in rocks of metamorphic character, having originated by metamorphism of silica-poor source rocks or, more commonly, by metasomatic replacement of rocks such as marbles and gneisses (as near Bancroft, Ontario, Canada). Cancrinite occurs principally with, and as a replacement of, nepheline in plutonic rocks, having been derived by the reaction of preexisting nepheline with $CaCO_3$-containing solutions. Sodalite also occurs principally in association with nepheline in syenites and related silica-poor intrusive igneous rocks, but it may also occur in equivalent extrusive rocks. Both haüyne and nosean are found almost exclusively in extrusive, silica-deficient rocks such as phonolites. Leucite occurs in potassium-rich basic lavas, frequently in large phenocrysts having the form of trapezohedrons. Pseudoleucite is a mixture of feldspar and nepheline which has the crystal form of preexisting leucite; it occurs in both extrusive and intrusive rocks. Unlike most other members of the feldspathoid group, lazurite occurs in metamorphic rocks, specifically, metamorphosed impure limestones. Lapis lazuli is the name of a blue lazurite-rich rock which is used as a gem material. *See* CANCRINITE; LAZURITE; NEPHELINITE; SILICATE MINERALS.

Donald R. Peacor

Feline infectious peritonitis

A fatal disease of both domestic and exotic cats (particularly cheetahs) caused by feline infectious peritonitis virus, a member of the Coronaviridae family. There are multiple strains of the virus which vary in virulence. Feline infectious peritonitis virus is closely related morphologically, genetically, and antigenically to other members of the Coronaviridae and arises as mutants from feline enteric coronaviruses, which infect cats but generally induce very mild or inapparent gastroenteritis. All coronaviruses are single-stranded ribonucleic acid (RNA) viruses with poor or no error correction during replication, resulting in relatively high mutation rates.

Epidemiology and transmission. Feline coronaviruses are contracted primarily during exposure to infectious cat feces in the environment, but also via ingestion or inhalation during cat-to-cat contact. Feline infectious peritonitis virus is relatively labile once outside the cat's body but may be able to survive for as long as 7 weeks if protected from heat, light, and desiccation. It is readily inactivated by most disinfectants.

Cases of feline infectious peritonitis occur either sporadically or in outbreaks with 50–70% mortality in homes with more than one cat. Most sporadic cases occur when a cat is infected with a feline enteric coronavirus which then mutates within the individual cat into the pathogenic feline infectious peritonitis phenotype. Cats with feline infectious peritonitis may shed either benign enteric coronavirus, potentially fatal feline infectious peritonitis virus, or neither. Many cats do not succumb to feline infectious peritonitis even after prolonged extensive exposure to infected cats.

In contrast to the sporadic transmission of feline infectious peritonitis, the transmission of the parental feline enteric coronavirus is frequent and is difficult to detect clinically. Prevalence of infections with enteric coronavirus is related to cattery size and density: in homes with five or more cats, approximately 100% of the cats have been exposed. Following exposure (often as kittens), cats may periodically shed enteric coronavirus in feces for weeks or a few months. After recovery, cats are not immune to the virus. Thus, they are likely to become rapidly reinfected and to shed again. Approximately 5% of kittens become chronic shedders after infection. Very few of the infected cats will show any clinical signs of enteritis; however, 2–5% of the cats may experience mutation to the feline infectious peritonitis phenotype and develop clinical signs of the virus. The likelihood of the mutant form developing and not being rapidly eliminated by the cat's immune system depends upon whether the cat is immunosuppressed (as by concurrent feline leukemia or feline immunodeficiency virus infection) and overall viral load. It is not known exactly at what point following infection with enteric coronavirus that the mutation occurs. The peak age distribution for development of feline infectious peritonitis is 6 months to 3 years of age.

Pathogenesis. In contrast to feline enteric coronavirus, whose replication is largely restricted to intestinal epithelial cells, feline infectious peritonitis virus targets feline macrophages (white blood cells). The virus can persist and replicate within the macrophages and spreads systemically through the bloodstream and into tissues throughout the body. The basis for development of clinical disease associated with feline infectious peritonitis depends upon a balance between antibodies and cell-mediated immunity (comprising T lymphocytes, natural killer cells, macrophages, and other innate immunity). The presence of antibodies to the virus enhances the access of the virus into the macrophage and may produce accelerated disease, although this phenomenon appears to be restricted to laboratory strains and infection protocols. Virus-antibody complexes are also deposited in the walls of small blood vessels, where they activate the complement system, attracting other blood vessels and initiating the profound inflammatory lesions that characterize the disease.

Clinical signs. Signs of infection with feline infectious peritonitis virus depend upon the severity of infection, the relative ability of the immune system to minimize some of the lesions, and the organ systems affected. If antibody-virus complexes are diffusely distributed along small blood vessels in the abdomen, the resulting immune-mediated sequelae may produce vasculitis which results in protein and fluid leakage out of the blood vessels and into the abdominal cavity (or less frequently thoracic and scrotal spaces). This form of disease is called wet feline infectious peritonitis and occurs in cats with overwhelming infection, with poor immunity, or during late stages of other forms of the virus. In contrast, if a cat has a moderately competent (but ultimately ineffectual) immune response to the virus infection, granulomas may arise in infected tissues which attempt to localize virus with an influx of uninfected macrophages, neutrophils, and lymphocytes. This dry form commonly affects kidneys, liver, lymph nodes, mesentery, diaphragm, the outer surface of the intestine, and the neurological system.

Cats with either wet or dry feline infectious peritonitis may exhibit weight loss, fever (often cyclical), and lethargy. Disease associated with a particular organ system reflects impairment or failure of that system, such as vomiting associated with liver disease, and seizures and ocular opacity associated with neurological or ocular feline infectious peritonitis. Wet or effusive feline infectious peritonitis is characterized by abdominal swelling, jaundice, and typically difficulty in breathing. Signs of upper respiratory disease or diarrhea are not directly associated with the virus. Suspicion of feline infectious peritonitis is based upon consistent history and signalment (typically cats originally from multiple-cat households which are 1–3 years old) and presence of consistent clinical signs. Suggestive laboratory testing includes elevated serum globulins, any positive coronavirus titer (which will not, however, discriminate between feline infectious peritonitis virus and feline enteric coronavirus), and mild elevations of neutrophils in blood and mild reductions in lymphocytes. The only accurate testing is biopsy of lesions or necropsy.

Treatment and control. As with most viral infections, there is no specific antiviral drug of proven efficacy in the treatment of feline infectious peritonitis. Clinical management rests upon palliative treatment of the specific signs, exhibited by each cat, and upon antibiotics, when indicated, to reduce secondary bacterial infections. The most important therapeutic approach involves the administration of immunosuppressive doses of corticosteroids to reduce the cat's immune response to the virus.

Attempts at prevention are frustratingly difficult, but include vaccination, testing and removing infected cats, and early weaning of kittens from queens. No one method is completely effective at eliminating enteric coronavirus or preventing feline infectious peritonitis. The best protection is by reducing group size to five or less. *See* ANIMAL VIRUS; VIRUS. Janet E. Foley

Bibliography. C. E. Greene, *Infectious Disease of the Dog and Cat*, Saunders, 1998; N. C. Pedersen, *Feline Infectious Diseases*, American Veterinary Publications, 1988.

Feline leukemia

A type of cancer caused by the feline leukemia virus, a retrovirus which affects only a small percentage of freely roaming or domestic cats. The feline leukemia virus is genetically and morphologically similar to murine leukemia virus, from which it presumably evolved several million years ago.

Epidemiology and transmission. About 1–5% of healthy-appearing wild or freely roaming domestic cats have lifelong (persistent) infections. These carrier cats shed the virus in urine, feces, and saliva. The principal route of infection is oral. Transmission in the uterus, transmission from mother to kitten through nursing, and intraspecies biting are of secondary importance.

The infection rate among freely roaming cats rises progressively from 6 months onward, and by 3–5 years of age a majority have been exposed. However, infections occurring in nature are usually inapparent or mild, and 95% of such cats recover without any signs of illness. The remainder become infected for life and shed the virus for months or years before they finally succumb. Mortality due to a feline leukemia virus infection occurs mainly among persistently infected cats and at a rate of around 50% per year. Almost 90% of these infected cats die within 3 years.

Although feline leukemia virus is of minor importance among freely roaming cats, the mortality can be 30–50% among certain high-risk indoor cat populations. Populations at risk include breeding catteries, shelters, managed wild cat colonies, pet stores, and homes with a large number of pet cats. The infection is more severe in these instances because indoor cats often share the same litter and food and water dishes, and frequently groom each other. Emotional stresses among large groups of confined cats further lower their resistance. Kittens are much less resistant than adults to persistent infection, making breeding groups of indoor animals particularly susceptible.

The typical cat infected with feline leukemia virus is an unvaccinated, nonpurebred cat 1–7 years of age that has been allowed to roam freely or has been inadvertently housed with an infected animal.

About one-half of cats infected with feline leukemia virus develop cancers (leukemias) arising from cells within the bone marrow, thymus gland, lymph node, or spleen. Another one-fourth die from anemia, and an equal number are affected by a type of acquired immunodeficiency syndrome with complicating opportunistic infections. A smaller proportion of cats infected with feline leukemia virus develop neurologic or ocular disorders, or other rare types of cancers.

Treatment. There is no treatment that eliminates the virus. Once the illness appears, a medical examination will determine precisely which feline leukemia virus-related disorder has occurred. Supportive or symptomatic treatment may prolong life for weeks or months, depending on the particular disease manifestation.

Prevention. All cats should be tested for the presence of the virus prior to putting them in contact with feline leukemia virus–free animals. Healthy-appearing or ill infected cats should not be in intimate contact with noninfected cats, even if the latter have been vaccinated. If they have to be kept in the same premises, they should be isolated from other cats and their litter and food and water dishes kept separate. The virus is very unstable once it is shed from the body; therefore there is very little transmission by contaminated clothing or hands.

Inactivated and genetically engineered subunit feline leukemia virus vaccines are available and should be administered annually. Because they are not totally effective in preventing infection, feline leukemia virus vaccines should not be considered a substitute for testing, elimination, and quarantine procedures. Vaccination is not essential for virus-free cats that are kept strictly confined indoors, providing precautions are taken to prevent contact with infected animals. *See* LEUKEMIA; RETROVIRUS. Niels C. Pedersen

Bibliography. J. A. Levy (ed.), *The Retroviridae*, Vol. 2, 1993; N. C. Pedersen, *Feline Infectious Diseases*, 1988.

Feline panleukopenia

An acute virus infection of cats, also called feline viral enteritis and (erroneously) feline distemper. The virus infects all members of the cat family (Felidae) as well as some mink, ferrets, and skunks (Mustelidae); raccoons and coatimundi (Procyonidae); and the binturong (Viverridae). Panleukopenia is the most important infectious disease of cats.

Infectious agent. Feline panleukopenia virus is classified as a parvovirus, and is one of the smallest known viruses. It is antigenically identical to the mink enteritis virus, and only minor antigenic differences exist between feline panleukopenia virus and canine parvovirus. It is believed that canine parvovirus originated as a mutation from feline panleukopenia virus. *See* ANIMAL VIRUS.

Epidemiology. This disease occurs worldwide, and nearly all cats are exposed by their first year because the virus is stable and ubiquitous; the disease is therefore rarely seen in older cats. Infected cats shed large amounts of virus in their secretions, vomit, and feces. The virus is highly resistant to inactivation and can remain infectious for months or years. It can easily by carried on shoes, clothes, feed bags, and toys. In addition, recovered cats may shed the virus for some time.

Pathogenesis. In order for parvoviruses to replicate, the cells they infect must be in a replicative stage. Consequently, only tissues with high cell proliferation rates are significantly affected by this virus group. These include leukocytes, bone marrow, intestinal crypt cells (cells that produce absorptive cells of intestinal villi) and many additional cell types in the developing fetus or neonate. Thus, the nature of the disease caused by feline panleukopenia virus depends upon the cat's age or pregnancy status. The typical form of the disease occurs in cats 2 weeks of age and older, which are exposed by nose or mouth. The virus attaches to and replicates in lymphoid cells lining the pharynx and adjacent tissues, and these cells carry the virus internally. A generalized infection occurs when new lymphoid cells are infected and the virus is released into the blood, infecting leukocytes, bone marrow, lymphoid cells of the thymus, lymph nodes, spleen, and follicles in the intestine. The virus then infects the proliferative crypt cells of the intestinal mucosa. Cells infected by feline panleukopenia virus die when the virus replicates, accounting for the clinical signs observed with panleukopenia. First, there is a precipitous drop in leucocyte counts 2–4 days after exposure. (The name panleukopenia, meaning absence of leukocytes in the blood, relates to this clinical feature.) Twenty-four to forty-eight hours later, destruction of the intestinal crypt cells results in the shortening and disruption of villi, causing diarrhea, vomiting, and loss of body fluids. Injury to the intestinal mucosa may allow bacteria or their toxins to enter the body. This, along with the destruction of the disease-fighting white blood cells, predisposes the cat to blood poisoning. Without treatment, this disease is often fatal.

The rapidly developing cat fetus may become infected and die, resulting in complete fetal resorption or delivery of a mummified or stillborn fetus. A unique syndrome, characterized by underdevelopment of the cerebellum, can occur in kittens infected by feline panleukopenia virus during the last 2 weeks of gestation or the first 2 weeks of life. The kitten may appear completely healthy, but it lacks muscle coordination because of destruction of parts of the cerebellum.

Diagnosis. Perhaps the majority of cats infected with feline panleukopenia virus do not become ill. However, the disease is severe and life threatening in 20–50% of cases. The cat is depressed and may refuse food or water; vomiting and diarrhea are common, resulting in severe dehydration. The cat may have a fever or a subnormal temperature. A low white blood cell count confirms the diagnosis as panleukopenia. Diagnosis can be made by autopsy and evidence of the destruction of the intestinal crypts and villus shortening.

Immunity and prevention. Both modified-live-virus and inactivated vaccines that are highly effective and safe are available for the prevention of panleukopenia. Only inactivated vaccines should be given to pregnant or immunosuppressed cats. The major obstacle to successful immunization of kittens is the presence of antibodies in the mother's milk. These antibodies block vaccination, and may last for the first 6–14 weeks of life, depending of the immune levels of the mother. For this reason, a series of two or three vaccinations, starting at 8 weeks of age and spaced 2–3 weeks apart, is given to kittens to ensure successful vaccination; a booster is often given once a year. If the immune status of a cat is unknown, and the cat will be in an environment with a high probability of exposure to feline panleukopenia virus (such as a cat show or a kennel), immune serum can be administered. This will protect the cat for 2–4 weeks, but the cat will have to be vaccinated later for permanent immunity.

Premises contaminated by feline panleukopenia virus are extremely difficult to disinfect because it is resistant to most of the common disinfectants; chlorine bleach, formaldehyde, or a certain quaternary ammonium disinfectant will destroy the virus. Because carpets, furniture, or other easily damaged objects in a home cannot be satisfactorily disinfected, a cat should be successfully immunized before being introduced to premises where a panleukopenia-infected cat previously lived. Jack H. Carlson

Bibliography. F. Fenner et al. (eds.), *Veterinary Virology*, 1993; B. N. Fields et al. (eds.), *Virology*, 3d ed., 1996; C. E. Greene (ed.), *Infectious Diseases of the Dog and Cat*, 2d ed., 1998.

Fennel

Forniculum vulgare, a culinary herb of the parsley family (Apiaceae). It is grown for the dried, ripe fruits or seeds which are used in bread, pickles, liqueurs, and meat sauces and dishes. Although similar in odor to anise, fennel seed can be distinguished by its warm, sweet character. *See* APIALES.

Fennel is one of approximately five species in the genus *Forniculum*. The plant is an erect, branching, short-lived perennial growing 3–5 ft (1–1.8 m) high. The leaves are three- to four-pinnately compound, threadlike, and 1.5 in. (3.7 cm) long. Common giant fennel (*Ferula communis*), a close relative, grows twice as large (6–12 ft or 2–4 m), but has no culinary use.

The three fennel varieties of commercial importance are *vulgare*, *dulce*, and *piperitum*. The variety *vulgare* is grown for its seed and the essential oil, obtained by steam distillation and known as bitter fennel oil. The variety *dulce* (finochio or Florence fennel) is grown for four products; seed and leaf for culinary use, the enlarged leaf base for a vegetable, and for the essential oil from the seeds (sweet fennel oil). The young stems of Italian fennel, variety *piperitum*, are used for flavoring in salads.

Fennel is native to southern Europe and the Mediterranean region and is presently cultivated as an annual or biennial plant in southern and eastern Europe, India, Argentina, China, and Pakistan. The plants do best in a mild climate and on limey, well-drained soils with plenty of sun. Sown in early spring at 8–9 lb/acre (9–10 kg/ha) with rows spaced 2–3 ft (0.6–0.9 m) apart, the fennel plants are thinned to 8–12 in. (20–30 cm) apart when established. The ripe seeds are greenish gray and fairly hard. When the plants are harvested, yields are low the first season but increase to 700–1000 lb/acre (800–1100 kg/ha) by the second year. Fennel should be located away from dill to avoid the resulting flavor effects from cross-pollination. Aphids and mildew can cause damage to fennel.

Though it is primarily a culinary herb, commercial interest in fennel also lies in its essential oil, bitter or sweet, for cooking, medicine, perfumery, and as a masking agent for insecticides and room sprays. *See* SPICE AND FLAVORING.

Seth Kirby

Fermentation

Decomposition of foodstuffs generally accompanied by the evolution of gas. The best-known example is alcoholic fermentation, in which sugar is converted into alcohol and carbon dioxide. This conversion, described by the equation below, was established

$$\underset{\text{Sugar}}{C_6H_{12}O_6} \rightarrow \underset{\substack{\text{Carbon}\\ \text{dioxide}}}{2CO_2} + \underset{\text{Alcohol}}{2C_2H_5OH}$$

by J. L. Gay-Lussac in 1815. *See* DISTILLED SPIRITS.

Before 1800 the association of yeast or leaven with fermentation had been noted, but the nature of these agents was not understood. Experiments of C. Cagniard-Latour, of F. T. Kützing, and of T Schwann in 1837 indicated that yeast is a living organism and is the cause of fermentation. This view was opposed by such leading chemists as J. von Liebig and F. Wöhler, who sought a chemical rather than a biological explanation of the process. The biological concept became generally accepted following the work of Louis Pasteur, who concluded that fermentation is a physiological counterpart of oxidation, and permits organisms to live and grow in the absence of air (anaerobically). This linked fermentation and putrefaction as comparable processes; both represent decompositions of organic matter brought about by microorganisms in the absence of air. The difference is determined by the nature of the decomposable material; sugary substances generally yield products with pleasant odor and taste (fermentation), whereas proteins give rise to evil-smelling products (putrefaction). *See* YEAST.

Pasteur also discovered the lactic acid and butyric acid fermentations, and from his experiments concluded that each kind of fermentation was caused by a specific microbe. Later work supported this idea to a large extent, and considerably increased the number of specific fermentations.

During fermentation organic matter is decomposed in the absence of air (oxygen); hence, there is always an accumulation of reduction products, or incomplete oxidation products. Some of these products (for example, alcohol and lactic acid) are of importance to society, and fermentation has therefore been used for their manufacture on an industrial scale. There are also many microbiological processes that go on in the presence of air while yielding incomplete oxidation products. Good examples are the formation of acetic acid (vinegar) from alcohol by vinegar bacteria, and of citric acid from sugar by certain molds (for example, *Aspergillus niger*). These microbial processes, too, have gained industrial importance, and are often referred to as fermentations, even though they do not conform to Pasteur's concept of fermentation as a decomposition in the absence of air. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; CITRIC ACID; INDUSTRIAL MICROBIOLOGY; VINEGAR.

Cornelis B. Van Niel

Bibliography. D. R. Berry (ed.), *Physiology of Industrial Fungi*, 1988; R. Dubos, *Louis Pasteur, Free Lance of Science*, 1960, reprint 1986; B. M. McNeil and L. M. Harvey (eds.), *Fermentation: A Practical Approach*, 1990; O. P. Ward, *Fermentation Biotechnology*, 1989.

Fermi-Dirac statistics

The statistical description of particles or systems of particles that satisfy the Pauli exclusion principle. This description was first given by E. Fermi, who applied the Pauli exclusion principle to the translational energy levels of a system of electrons. It was later shown by P. A. M. Dirac that this form of statistics is also obtained when the total wave function of the system is antisymmetrical. *See* EXCLUSION PRINCIPLE; NONRELATIVISTIC QUANTUM THEORY.

Distribution function. Such a system is described by a set of occupation numbers $\{n_i\}$ which specify the number of particles in energy levels ϵ_i. It is important to keep in mind that ϵ_i represents a finite range of energies, which in general contains a number, say g_i, of nondegenerate quantum states. In the Fermi statistics, at most one particle is allowed in a nondegenerate state. (If spin is taken into account, two particles may be contained in such a state.) This is simply a restatement of the Pauli exclusion principle, and means that $n_i \leqq g_i$. The probability of having a set $\{n_i\}$ distributed over the levels ϵ_i, which contain g_i nondegenerate levels, is described by Eq. (1),

$$W = \prod_i \frac{g_i!}{(g_i - n_i)!n_i!} \tag{1}$$

which gives just the number of ways that n_i can be picked out of g_i, which is intuitively what one expects for such a probability. In Boltzmann statistics this same probability is given by Eq. (2). The equilibrium state which actually exists is the set of n's that makes W a maximum, under the auxiliary conditions given in Eqs. (3). These conditions express the fact that the total energy E and the total number of particles N are given. *See* BOLTZMANN STATISTICS.

$$W' = \prod_i \frac{g_i^{ni}}{n_i!} \tag{2}$$

$$\Sigma n_i = N \tag{3a}$$

$$\Sigma n_i \epsilon_i = E \tag{3b}$$

Equation (4) holds for this most probable distribution. Here A and β are parameters, to be determined from Eq. (3); in fact, $\beta = 1/kT$, where k is Boltzmann's constant and T is the absolute temperature. When the 1 in the denominator may be neglected, Eq. (4) goes over into the Boltzmann distribution; this provides a procedure for identifying β. It is known that in classical statistics the Boltzmann distribution may be obtained if specific assumptions are made as to the number of collisions taking place. It is there assumed that the number of collisions per second in which molecules with velocities in cells i and j in phase space produce molecules with velocities in cells k and l is given by Eq. (5). Here $a^{kl}{}_{ij}$ is a geometrical factor. This leads to the Boltzmann distribution. The Fermi distribution, Eq. (4), may be obtained if instead of Eq. (5) one assumes Eq. (6) for the number of collisions. One observes from Eq. (6) the interesting quantum theoretical feature that the probability for a collision depends on the occupation numbers of the states into which the colliding particles will go. In particular, if these final states are filled up ($n_k = g_k$), no collision with that state as a final state can occur.

$$n_i = \frac{g_i}{\dfrac{1}{A}\epsilon^{\beta\epsilon_i} + 1} \tag{4}$$

$$A_{ij}^{kl} = n_i n_j a_{ij}^{kl} \tag{5}$$

$$A_{ij}^{kl} = a_{ij}^{kl} n_i n_j \left(\frac{g_k - n_k}{g_k}\right)\left(\frac{g_l - n_l}{g_l}\right) \tag{6}$$

The distribution f_i is often used; it is defined by Eq. (7).

$$n_i = f_i g_i \tag{7}$$

Applications. For a system of N electrons, each of mass m in a volume V, Eq. (4) may be written as Eq. (8), where h is Planck's constant. Equations (3*a*) and (3*b*) may now be transformed into integrals, yielding Eqs. (9). Here $\lambda = (h^2/2\pi mkT)^{1/2}$ is the thermal de Broglie wavelength, and $U_\rho(A)$ is the Sommerfeld integral defined by Eqs. (10), where $u = mv^2/2kt$ and Γ is the usual Γ-function. Very often one writes instead of A in the Fermi distribution a quantity μ defined by Eq. (11). It may be shown (for instance, by going to the classical limit) that μ is the chemical potential. It may be seen that if ϵ is large $(1/kT)(\epsilon - \mu) \gg 1$; hence the Fermi distribution goes over into a Maxwell-Boltzmann distribution. It is easy to verify that this inequality may be transcribed so as to state expression (12). Physically this is reasonable because expression (12) says that classical conditions pertain when the volume per particle is much larger than the volume associated with the de Broglie wavelength of a particle. For example, $V/N\lambda^3$ is about 7.5 for helium gas (He^3) at 4 K (−452°F) and 1 atm (10^2 kilopascals). This indicates that classical statistics may perhaps be applied, although quantum effects surely play a role. For electrons in a metal at 300 K (80°F), $V/N\lambda^3$ has the value 10^{-4}, showing that classical statistics fail altogether for electrons in metals. When the classical distribution fails, a degenerate Fermi distribution results. Numerically, if $A \gg 1$, a degenerate Fermi distribution results; if $A \ll 1$, the classical results are again obtained. For example, Eqs. (9*a*), (9*b*), and (10*b*) show that $E = {}^3/_2 NkT$ and the specific heat should be ${}^3/_2 R = {}^3/_2 Nk$. However, for an electron gas this does not apply since such a system is degenerate for normal temperatures and $A \gg 1$. A somewhat lengthy calculation yields the result that, in the case $A \gg 1$, the contribution to the specific heat

$$f\left(v_x v_y v_z\right)\, dv_x\, dv_y\, dv_z = 2\left(\frac{m}{h}\right)^3 V \frac{dv_x\, dv_y\, dv_z}{\dfrac{1}{A}e^{mv^2/2kT} + 1} \tag{8}$$

$$\frac{N}{V}\frac{\lambda^3}{2} = U_{1/2}(A) \tag{9a}$$

$$\frac{E}{3kT}\frac{\lambda^3}{V} = U_{3/2}(A) \tag{9b}$$

$$U_\rho(A) = \frac{1}{\Gamma(\rho+1)}\int_0^\infty \frac{u^\rho\, du}{\dfrac{1}{A}e^u + 1} \tag{10a}$$

$$U_\rho(A) \cong A \quad (A \ll 1) \tag{10b}$$

$$U_\rho(A) \cong \frac{(\ln A)^{\rho+1}}{\Gamma(\rho+1)} \quad (A \gg 1) \tag{10c}$$

$$\mu = kT \ln A \qquad A = e^{\mu/kT} \tag{11}$$

$$\frac{V}{N} \gg \lambda^3 \tag{12}$$

is negligible. This resolves an old paradox, for, according to the classical equipartition law, the electronic specific heat C should be $^3/_2Nk$, whereas in reality it is very small; in fact, $C = \gamma T$, where γ is a very small constant. This is a consequence of the fact that an electron gas at normal temperature is a degenerate Fermi gas. The electrical resistance of a metal can be understood on a classical picture, but the lack of a specific heat is a pure quantum effect. *See* BOSE-EINSTEIN STATISTICS; FREE-ELECTRON THEORY OF METALS; GAMMA FUNCTION; KINETIC THEORY OF MATTER; QUANTUM STATISTICS; SPECIFIC HEAT OF SOLIDS; STATISTICAL MECHANICS. Max Dresden

Fermi surface

The surface in the electronic wavenumber space of a metal that separates occupied from empty states. Every possible state of an electron in a metal can be specified by the three components of its momentum, or wavenumber. The name derives from the fact that half-integral spin particles, such as electrons, obey Fermi-Dirac statistics and at zero temperature fill all levels up to a maximum energy called the Fermi energy, with the remaining states empty. *See* FERMI-DIRAC STATISTICS.

The fact that such a surface exists for any metal, and the first direct experimental determination of a Fermi surface (for copper) in 1957, were central to the development of the theory of metals. The Fermi surface is a geometrical shape which could be determined by experiment and calculated theoretically, providing the needed test of the quantum theory of metals. Neither step could have been taken much earlier. High-purity materials were needed for the experimental studies, and high-speed digital computers for the theoretical determinations. Before 1957, many scientists believed that the one-electron approximation was hopelessly naive. Soon afterward the Fermi Surface Conference (Cooperstown, New York, 1960) considered the many Fermi surfaces that had by that time been determined experimentally.

Nearly free electrons. An early surprise arising from the known Fermi surfaces was that many of the shapes were close to what would be expected if the electrons interacted only weakly with the crystalline lattice. The long-standing free-electron theory of metals was based upon this assumption, but most physicists regarded it as a serious oversimplification. *See* FREE-ELECTRON THEORY OF METALS.

The momentum p of a free electron is related to the wavelength λ of the electronic wave by Eq. (1),

$$p = \frac{2\pi\hbar}{\lambda} \qquad (1)$$

where $\hbar$ is Planck's constant divided by 2π. The ratio $2\pi/\lambda$, taken as a vector in the direction of the momentum, is called the wavenumber $\mathbf{k}$. If the electron did not interact with the metallic lattice, the energy would not depend upon the direction of $\mathbf{k}$, and all constant-energy surfaces, including the Fermi surface, would be spherical. This is illustrated in **Fig. 1**, with the radius $\mathbf{k}_F$ related to the number N of electrons per unit volume by Eq. (2).

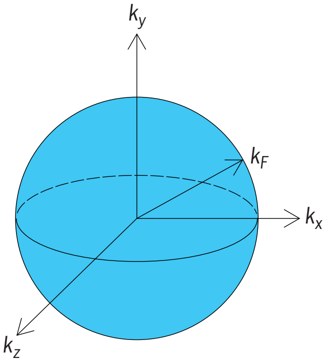


Fig. 1. Free-electron Fermi surface in the space defined by the components of wavenumbers of the electrons.

$$\mathbf{k}_F{}^3 = 3\pi^2 N \qquad (2)$$

The Fermi surface of copper was found to be distorted (**Fig. 2**) but was still a recognizable deformation of a sphere. The polyhedron surrounding the Fermi surface in the figure is called the Brillouin zone. It consists of Bragg-reflection planes, the planes made up of the wavenumbers for which an electron can be diffracted by the periodic crystalline lattice. The square faces, for example, correspond to components of the wavenumber along one coordinate axis equal to $2\pi/a$, where a is the cube edge for the copper lattice. For copper the electrons interact with the lattice so strongly that when the electron has a wavenumber near to the diffraction condition, its motion and energy are affected and the Fermi surface is correspondingly distorted. The Fermi surfaces of sodium and potassium, which also have one conducting electron per atom, are very close to the sphere shown in Fig. 1. These alkali

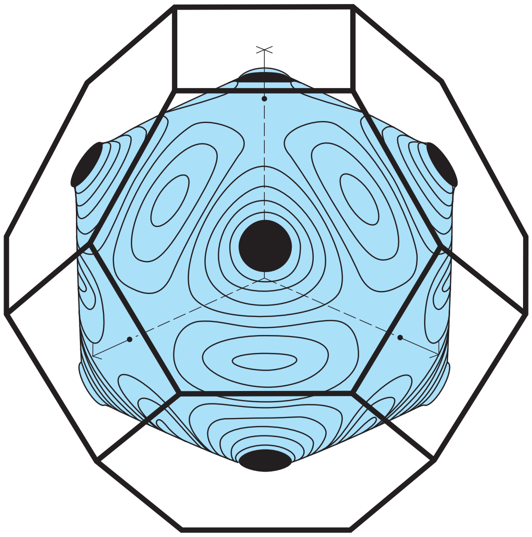


Fig. 2. Fermi surface of copper, as determined in 1957; two shapes were found to be consistent with the original data, and the other, slightly more deformed version turned out to be correct.

metals are therefore more nearly free-electron-like. *See* BRILLOUIN ZONE.

Metals with more than one electron per atom, such as calcium and aluminum, have higher electron densities and correspondingly larger Fermi wavenumbers k_F; the Fermi sphere overlaps the Bragg reflection planes. Again the interaction with the lattice is weak and deforms the Fermi surface only slightly. However, electrons accelerated by a magnetic field and with wavenumbers in a Bragg plane are diffracted to other parts of the Fermi sphere. Thus their motion in this wavenumber space is discontinuous, and the corresponding motion of the electron in real space shows sharp turns between smooth circular sections of orbit. It is possible to describe these complex motions in a more continuous fashion by translating various sections of the Fermi sphere back into the original Brillouin zone (Fig. 2), so that the motion of the wavenumber is continuous across these more complicated Fermi surfaces. Such constructions are shown in **Fig. 3** for calcium (valence 2) and aluminum (valence 3). The motion of any electron in a uniform magnetic field is along the intersection of one of these surfaces and a plane perpendicular to the magnetic field. The electron orbit in the real crystal is easily seen to be of exactly the same shape, though rotated by 90°. The size of the real orbit, for attainable magnetic fields, is very much larger than the lattice distance as the electron successively accelerates and diffracts. In the real metal, the Fermi surface and the corresponding orbits are deformed and rounded at the edges but are still close to the ideal surfaces shown.

The weakness of the interaction between electrons and lattice was a surprise, since the potential that describes this interaction is large. It turns out, however, that the electrons move so rapidly past the atoms that the effective potential, called a pseudopotential, is weak. It was the understanding of the nearly free-electron Fermi surfaces which suggested the weak pseudopotential and the possibility of simple theories of metals treating the pseudopotential as a small perturbation in the electron states.

In transition metals there are electrons arising from atomic *d* levels, in addition to the free electrons, and the corresponding Fermi surfaces are more complex than those of the nearly free-electron metals. However, the Fermi surfaces exist and have been determined experimentally for essentially all elemental metals.

Determination from experiments. The motion of the electrons in a magnetic field discussed above provides the key to experimentally determining the Fermi surface shapes. The simplest method conceptually derives from ultrasonic attenuation. Sound waves of known wavelength pass through the metal and a magnetic field is adjusted, yielding fluctuations in the attenuation as the orbit sizes match the sound wavelength. This measures the diameter of the orbit and Fermi surface. The most precise method uses the de Haas–van Alphen effect, based upon the quantization of the electronic orbits in a magnetic field. Fluctuations in the magnetic susceptibility give a direct measure of the cross-sectional areas of the Fermi surface. The method used in 1957 to determine the Fermi surface of copper did not use a magnetic field, but was based upon the anomalous skin effect, which determines the number of electrons moving parallel to the metal surface and from that deduces the curvature of the Fermi surface. *See* DE HAAS-VAN ALPHEN EFFECT; SKIN EFFECT (ELECTRICITY); ULTRASONICS.

Fermi surface determination has become an end in itself, with findings tabulated in the literature. There is intense interest in determining the Fermi

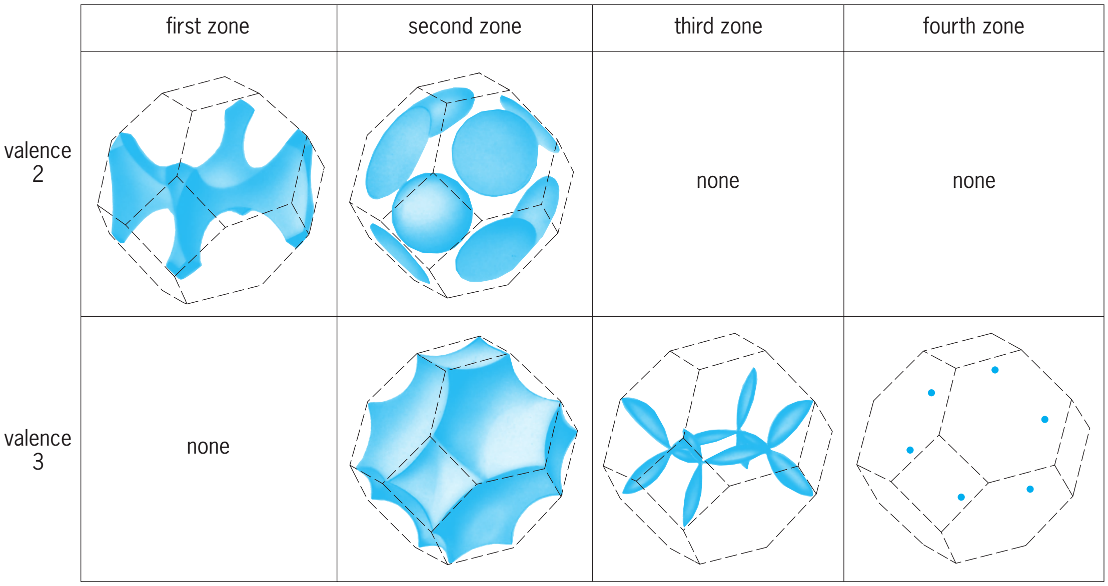


Fig. 3. Free-electron Fermi surfaces for face-centered cubic metals of valence 2 (calcium) and valence 3 (aluminum). (*After W. A. Harrison, Electronic structure of polyvalent metals, Phys. Rev., 118:1190–1208, 1960*)

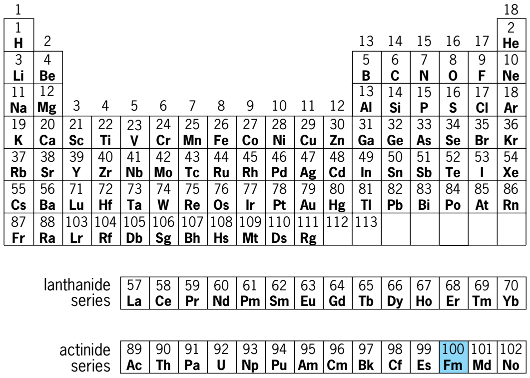
surface, if one exists, in the high-temperature superconductors, but no experimental determination has yet been convincing. *See* BAND THEORY OF SOLIDS; SOLID-STATE PHYSICS; SUPERCONDUCTIVITY.

Walter A. Harrison

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; A. P. Cracknell (ed.), *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, *New Series*, Group 3: *Crystal and Solid State Physics*, vol. 13c: *Electron States and Fermi Surfaces of Elements*, 1984; A. P. Cracknell and K. C. Wong, *The Fermi Surface*: *Its Concept*, *Determination*, *and Use in the Physics of Metals*, 1973; W. A. Harrison and M. B. Webb (eds.), *The Fermi Surface*, 1960; J. Kondo and A. Yoshimori, *Fermi Surface Effects*, 1988.

Fermium

A chemical element, Fm, atomic number 100, the eleventh element in the actinide series. Fermium does not occur in nature; its discovery and production have been accomplished by artificial nuclear transmutation of lighter elements. Radioactive isotopes of mass number 244–259 have been discovered. The total weight of fermium which has been synthesized is much less than one-millionth of a gram. *See* ACTINIDE ELEMENTS; PERIODIC TABLE; RADIOACTIVITY.



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Spontaneous fission is the major mode of decay for ^{244}Fm, ^{256}Fm, and ^{258}Fm. The longest-lived isotope is ^{257}Fm, which has a half-life of about 100 days. Fermium-258 decays by spontaneous fission with a half-life of 0.38 millisecond. This suggests the existence of an abnormality at this point in the nuclear periodic table. *See* NUCLEAR CHEMISTRY; NUCLEAR REACTION; TRANSURANIUM ELEMENTS.

Glenn T. Seaborg

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; J. J. Katz et al. (eds.), *The Chemistry of the Actinide and Transactinide Elements*, 3d ed., 2006; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Ferret

One of three carnivorous mammals, also known as polecats, in the family Mustelidae. The European polecat (*Mustela putorius*) inhabits Europe west of the Ural Mountains; the steppe polecat (*M. eversmanni*) is found in the steppe zone from Austria to Manchuria and Tibet; and the black-footed ferret (*M. nigripes*) inhabits the plains region of North America from Alberta and Saskatchewan to northeastern Arizona and Texas. The domestic ferret (*M. putorius furo*) is generally thought to be a descendant of one or both of the Old World species.

General morphology. Ferrets and polecats are weasellike and have long slender bodies, small rounded ears, and short legs. The limbs bear five digits with nonretractile curved claws. The European polecat is the darkest of the three species with a general coloration of dark brown to black with pale yellow underfur. The steppe polecat is generally straw yellow or pale brown with the chest, limbs, and terminal third of the tail being dark brown to black. There is a dark mask across the face. The black-footed ferret is generally yellow buff with the forehead, muzzle, and throat being nearly white. The top of the head and middle of the back are brown, while the face mask, feet, and tip of the tail are black. The domestic ferret is generally white or pale yellow in color. The dental formula is I 3/3, C 1/1, Pm 3/3, M 1/2 × 2 for a total of 34 teeth. The senses of hearing, smell, and sight are well developed. The head and body length of these three species ranges 205–500 mm (8–20 in.), and the tail length is 70–190 mm (2.7–7.5 in.). Weight ranges 205–2050 g (0.75–4.5 lb). Males are larger than females. *See* DENTITION.

Range and occurrence. The original range of the black-footed ferret coincided with the range of prairie dogs (*Cynomys*), its primary prey. The widespread use of sodium fluoroacetate (“1080”) to poison prairie dogs resulted in secondary poisoning of black-footed ferrets that fed on them. In addition, the large-scale eradication of prairie dogs from the range of the black-footed ferret greatly reduced their

Black-footed ferret, *Mustela nigripes*. (© *Joseph Dougherty/www.ecology.org*)

food supply. Today, the ferret is considered the most endangered mammal in North America. One of the last known populations of black-footed ferrets was extirpated in South Dakota during the early 1970s. The species was feared to be extinct until a small population was found in northwestern Wyoming in 1981. Several animals from the Wyoming population were caught for a captive breeding program before the wild population disappeared. Successes in the captive breeding program allowed reintroductions at a southcentral Wyoming site from 1991 until 1995, but disease and other factors caused that program to be suspended. Beginning in 1994, reintroduction efforts were begun in Montana and South Dakota. In 1996, Arizona became the recovery program's fourth reintroduction site. At present, there are no known nonintroduced wild populations. *See* PRAIRIE DOG.

Behavioral ecology and reproduction. Relatively little is known about the ecology of black-footed ferrets. They are primarily nocturnal and spend most of their time in underground prairie dog burrows which they use for shelter and travel. Adult ferrets are solitary except during the breeding season. Adults of the same sex are not known to inhabit the same prairie dog town, whereas adults of the opposite sex are found in the same town but not in the same burrow system. While prairie dogs appear to be the primary food, black-footed ferrets also prey on a variety of other small mammals. They kill their prey by attacking the neck and base of the skull. Potential predators of ferrets are badgers, coyotes, bobcats, rattlesnakes, eagles, hawks, and owls.

Ferrets breed in the spring with litters born after a gestation period of 42–45 days. Litter size ranges 1–6 with an average of 3.5 kits. Unlike many weasels, there does not seem to be delayed uterine implantation. Newborn young are altricial (immature and helpless) and have their eyes closed. Sexual maturity is attained by one year of age. Captives have lived to about 12 years of age.

The domestic ferret is usually tame and playful, and is found in captivity throughout the world. In some areas it is used to control rodents and to drive rabbits from their burrows. *See* CARNIVORA; WEASEL.

Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Ferricyanide and ferrocyanide

The common names for hexacyanoferrate(III) and hexacyanoferrate(II), respectively.

Ferricyanide. Hexacyanoferrate(III) is a compound containing the $[Fe(CN)_6]^{3-}$ complex ion. The oxidation state of iron is a tripositive ion [Fe^{3+} or Fe^{III}] and is low spin (spin paired), consistent with the strong field nature of the cyanide (CN^-) ligand. The complex $[Fe(CN)_6]^{3-}$ adopts an octahedral geometry (see **illus.**).

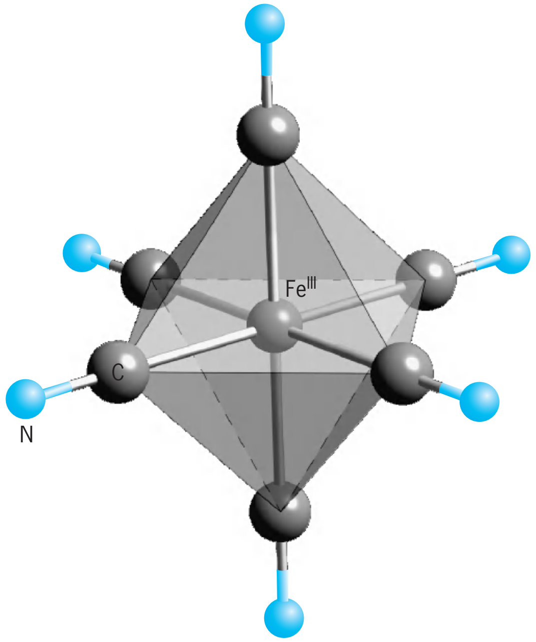


Diagram showing octahedral shape of ferricyanide.

The $[Fe(CN)_6]^{3-}$ ion is kinetically unstable, as it dissociates to give the free cyanide anion, CN^-. It is therefore toxic ($LD_{50} = 2690$ mg/kg versus 10 mg/kg for KCN). In contrast, the ferrocyanide ion $[Fe(CN)_6]^{4-}$ is stable and less toxic ($LD_{50} = 6400$ mg/kg). Substituted derivatives, such as the nitroprusside ion $[Fe(CN)_5NO]^{2-}$, are known. The redox potential for the aqueous reduction of ferricyanide to ferrocyanide is ($[Fe(CN)_6]^{3-} + e^- = [Fe(CN)_6]^{4-}$), $E^\circ = 0.36$ V (versus standard hydrogen electrode), but is quite solvent dependent.

The sodium ($Na_3[Fe(CN)_6]$) and potassium ($K_3[Fe(CN)_6]$) salts have been isolated as ruby-red crystals and are photosensitive. The potassium salt reacts with metallic silver to produce silver ferrocyanide, which is used in photographic processes. Addition of Fe^{II} ions to ferricyanide forms Prussian blue $Fe^{III}_4\,[Fe^{II}(CN)_6]_3 \cdot xH_2O$, $x \sim 15$. This pigment is used in blueprint materials and as a mild oxidizing agent in organic synthesis. *See* PRUSSIAN BLUE.

Ferrocyanide. Hexacyanoferrate(II) is a compound containing the complex ion $[Fe(CN)_6]^{4-}$. The oxidation state of iron is dipositive (Fe^{2+} or Fe^{II}) and is low spin (spin paired), consistent with the strong field nature of the cyanide (CN^-) ligand.

The $[Fe(CN)_6]^{4-}$ ion is very stable, with an octahedral geometry similar to that of $[Fe(CN)_6]^{3-}$. The free acid, hydrogen hexacyanoferrate(II) ($H_4[Fe(CN)_6]$), is soluble in water and is isolated as a white powder by precipitation of the ion in strongly acidic solutions. Salts of sodium ($Na_4[Fe(CN)_6]$) are prepared by heating an aqueous mixture of sodium cyanide (NaCN), iron(II) sulfate ($FeSO_4$), and other salts. In contrast to ferricyanide $[Fe(CN)_6]^{3-}$, ferrocyanide is substitutionally inert—that is, it is slow to liberate cyanide. Salts of ferrocyanide are yellow hydrates, which are insoluble in most organic solvents. Ferrocyanide is also used as an antidote for thallium poisoning.

Hexacyanoferrate(II) may be converted to hexacyanoferrate(III) (ferricyanide) by strong oxidizing agents such as peroxides and permanganate ion.

Ferrocyanide is commonly used as a reducing agent. Its also is used as an additive to make free-flowing table salt and as a U.S. FDA-approved colorant, as well as in the preparation of dyes and photographic fixatives and for stabilizing synthetic and natural latex foams.

Replacement of Fe^{II} or Fe^{III} ions with other metal ions leads to other Prussian blue–structured materials, which are magnets. Use of Cr^{III} in place of Fe^{II}, as well as V^{II} in place of Fe^{II}, leads to a family of magnets with ordering temperatures above room temperature. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES; CURIE TEMPERATURE; CYANIDE; IRON.

Joel S. Miller

Bibliography. K. R. Dunbar and R. A. Heintz, Chemistry of transition metal cyanide compounds: Modern perspectives, *Progr. Inorg. Chem.*, 45:283–392, 1997; J. S. Miller, 3-D-network structured cyanide-based magnets, *MRS Bull.*, 25(11):60–64, 2000; A. G. Sharp, *The Chemistry of Cyano Complexes of the Transition Metals*, Academic Press, New York, 1976.

Ferrimagnetic garnets

A class of ferrimagnetic oxide materials that have the garnet crystal structure. The crystal structure contains three types of nonequivalent sites in the oxygen lattice which are occupied by metal ions. The sites are referred to as 16(*a*), 24(*c*), and 24(*d*). A metal ion in a 16(*a*) site is surrounded by six oxygen ions which are at the corners of the octahedron. The metal ions in the 24(*d*) sites are found in nonregular tetrahedrons formed by oxygen ions, while the metal ions in the 24(*c*) sites have eight oxygen-ion nearest neighbors to form a triangular dodecahedron at the corners of a distorted cube. *See* CRYSTALLOGRAPHY; FERRIMAGNETISM; FERRITE.

The typical formula of a ferrimagnetic garnet is $X_3Fe_5O_{12}$, where the trivalent X ion is yttrium, or any of the rare-earth ions with an atomic number greater than 61. (The rare-earth ions below samarium apparently have radii too large to fit into the garnet structure.) However, the larger rare earths and lanthanum can substitute partially for some of the yttrium or smaller rare earths. The iron occupies both the 16(*a*) and 24(*d*) sites, while the yttrium, lanthanum, or rare-earth ion, or a combination of two or more of these ions, occupies the 24(*c*) sites. The 24 iron ions (per unit cell) on the (*d*) sites are strongly coupled, with an antiferromagnetic coupling, to the 16 iron ions in the (*a*) sites. If the 24 ions on the (*c*) sites have a magnetic moment, it is weakly aligned antiparallel to the moment of the ions on the (*d*) sites. A significant amount of divalent calcium ion can be substituted for rare-earth ions, if an equal amount of a quadravalent ion such as germanium is substituted for iron to maintain charge balance. *See* RARE-EARTH ELEMENTS.

Ferrimagnetic garnets are of great theoretical interest because they have a highly ordered structure and because they accommodate rare-earth ions, some of which have a small contribution to their magnetism due to the orbital motion of electrons, in addition to the magnetism due to electron spin. The first practical engineering interest in the ferrimagnetic garnets was due to the yttrium iron garnet, which is used in certain microwave ferrite devices because of its very narrow ferromagnetic resonance absorption line (less than 1 oersted or 80 A/m at 10,000 MHz). Development of magnetic bubbles for use in solid-state nonvolatile memory devices resulted in a large research effort on complex rare-earth iron garnets. In these garnets several different rare-earth ions are simultaneously incorporated in the structure in concentrations designed to optimize a number of physical and magnetic properties which influence the memory parameters. An example is the garnet $(Y_{1.73}Lu_{.22}Sm_{.20}Ca_{.85})(F_{4.15}Ge_{.85})O_{12}$ used in practical bubble memories with a storage density of 3×10^5 bits/cm^2. *See* COMPUTER STORAGE TECHNOLOGY; FERRITE DEVICES; GYRATOR.

Single crystals of ferrimagnetic garnets can be grown directly from a melt of the appropriate oxides, from a solution of the oxides dissolved in a PbO/B_2O_3 flux, or from a gas of the chlorides, which are volatile at high temperatures. Rare-earth garnets for magnetic bubble devices are grown from the flux onto the surface of crystallographically compatible nonmagnetic gadolinium gallium garnet in the form of single-crystal epitaxial films a few micrometers thick. Under the growth conditions used, these films have anisotropic magnetic properties, even though the crystal structure is cubic, as a result of growth- and stress-induced effects. *See* CRYSTAL GROWTH.

Morton E. Jones

Bibliography. J. Goldman, *Modern Ferrite Technology*, 1990; A. Paoletti (ed.), *Physics of Magnetic Garnets*, 1979; C. Srivastava and M. J. Patni, *Ferrites*, 1990; H. Watanabe (ed.), *Ferrites*, 1982.

Ferrimagnetism

A specific type of ordering in a system of magnetic moments or the magnetic behavior resulting from such order. In some magnetic materials the magnetic ions in a crystal unit cell may differ in their magnetic properties. This is clearly so when some of the ions are of different species. It is also true for similar ions occupying crystallographically inequivalent sites. Such ions differ in their interactions with other ions, because the dominant exchange interaction is mediated by the neighboring nonmagnetic ions. They also experience different crystal electric fields, and these affect the magnetic anisotropy of the ion. A collection of all the magnetic sites in a crystal with identical behavior is referred to as a magnetic sublattice. A material is said to exhibit ferrimagnetic order when, first, all moments on a given sublattice point in a single direction and, second, the resultant moments of the sublattices lie parallel or antiparallel to one another. The notion of such an order is due to L. Néel, who showed in 1948 that its existence would explain many of the properties of the magnetic ferrites. *See* FERRITE; FERROMAGNETISM.

At high temperatures all magnetic systems are disordered. As the temperature of a potentially ferrimagnetic system is lowered, there comes a point, the Curie temperature, at which all sublattices simultaneously acquire a moment and arrange themselves in a definite set of orientations. The Curie temperature and the temperature dependence of the sublattice moments depend in a complicated way on the magnetic properties of the individual ions and upon the interactions between them. In general, there is a net moment, the algebraic sum of the sublattice moments, just as for a normal ferromagnet. However, its variation with temperature rarely exhibits the very simple behavior of the normal ferromagnet. For example, in some materials, as the temperature is raised over a certain range, the magnetization may first decrease to zero and then increase again. Ferrimagnets can be expected, in their bulk properties, measured statically or at low frequencies, to resemble ferromagnets with unusual temperature characteristics. *See* CURIE TEMPERATURE.

The most versatile of ferrimagnetic systems are the rare-earth iron garnets. The garnet unit cell has three sets of inequivalent magnetic sites, differing in their coordination to neighboring oxygen ions. Two of the sets, with 24 and 16 sites respectively per unit cell, are each occupied by Fe^{3+} ions, and the corresponding sublattices orient antiparallel. The remaining set of 24 sites may be occupied by nonmagnetic ions, such as Y, or by a magnetic rare-earth ion. This sublattice, when magnetic, usually lies parallel to the 24-site Fe^{3+} lattice. Rare-earth ions of various species have widely different magnetic moments, exchange interactions, and crystalline anisotropies. By a suitable choice of rare-earth ions, it is possible to design ferrimagnetic systems with prescribed magnetizations and temperature behavior. *See* FERRIMAGNETIC GARNETS.

Laurence R. Walker

Bibliography. D. J. Craik (ed.), *Magnetic Oxides*, 1975; D. J. Craik, *Magnetism*, 1994; C. Srivastava and M. J. Patni, *Ferrites*, 1990; H. Watanabe (ed.), *Ferrites*, 1982.

Ferrite

Any of the class of magnetic oxides. Typically the ferrites have a crystal structure that has more than one type of site for the cations. Usually the magnetic moments of the metal ions on sites of one type are parallel to each other, and antiparallel to the moments on at least one site of another type. Thus ferrites exhibit ferrimagnetism, a term coined by L. Néel. *See* FERRIMAGNETISM.

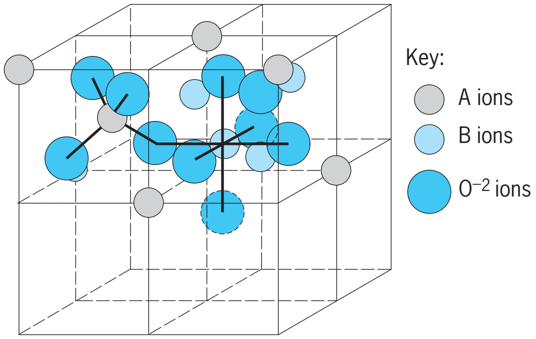


Fig. 1. Two octants of the spinel unit cell showing A ions on tetrahedral sites and B ions on octahedral sites. (*After E. P. Wohlfarth, ed., Ferromagnetic Materials, vol. 2, North-Holland, 1980*)

Commercial types. There are three important classes of commercial ferrites. One class has the spinel structure, with the general formula $M^{2+}Fe_2{}^{3+}O_4$, where M^{2+} is a divalent metal ion. In the unit cell of this structure, there are 8 tetrahedral and 16 octahedral sites for the locations (**Fig. 1**). The locations in the tetrahedral and octahedral sites are surrounded by four and six oxygen ions, respectively. The sites of each type constitute a sublattice. The moments of the two sublattices in the spinels are antiparallel, and the net moment is given by the difference in the moments of the sublattices. The magnetic moment of a spinel is then determined by the magnetic moment of M^{2+} and Fe^{3+} together with their distribution in the two sublattices. So-called linear ferrites used in inductors and transformers are made of Mn and Zn (for frequencies up to 1 MHz) and Ni and Zn (for frequencies greater than 1 MHz). MgMn ferrites are used in microwave devices such as isolators and circulators. Until the late 1970s, ferrites with square loop shapes held a dominant position as computer memory-core elements, but these gave way to semiconductors. *See* COMPUTER STORAGE TECHNOLOGY; GYRATOR; SEMICONDUCTOR.

The second class of commercially important ferrites have the garnet structure, with the formula $M_3{}^{3+}Fe_5{}^{3+}O_{12}$, where M^{3+} is a rare-earth or yttrium ion. The Fe^{3+} ions form two antiparallel sublattices composed of tetrahedral and octahedral sites. The M^{3+} ions form a third sublattice of dodecahedral sites in which each M^{3+} ion is surrounded by eight oxygen ions (**Fig. 2**). Yttrium-based garnets are used in microwave devices. Thin monocrystalline films of complex garnets have been developed for bubble domain memory devices. *See* FERRIMAGNETIC GARNETS.

The third class of ferrites has a hexagonal structure, of the $M^{2+}Fe_{12}{}^{3+}O_{19}$ magnetoplumbite type, where M^{2+} is usually Ba, Sr, or Pb. The Fe^{3+} ions occupy both tetrahedral and octahedral sites, in addition to a site surrounded by five oxygen ions. Because of their large magnetocrystalline anisotropy, the hexagonal ferrites develop high coercivity and are an important member of the permanent magnet family.

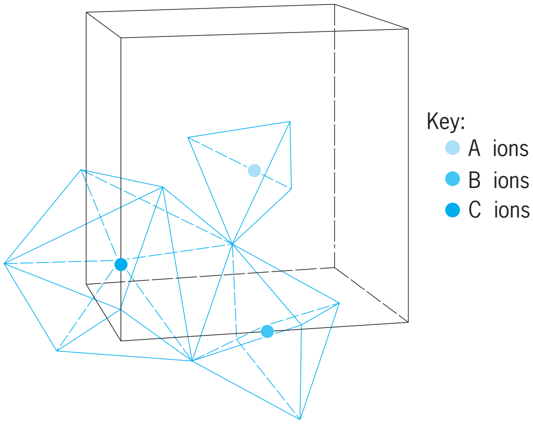


Fig. 2. One octant of the $M_3Fe_5O_{12}$ garnet unit cell showing A ions on tetrahedral sites, B ions on octahedral sites, and C ions on dodecahedral sites. (*After M. A. Gilleo and S. Geller, Magnetic and crystallographic properties of substituted yttrium-iron garnet, $3Y_2O_3 \cdot xM_2O_3 \cdot [5-x]\ Fe_2O_3$, Phys. Rev., 110:73–78, 1958*)

Another magnetic oxide, γ-Fe_2O_3, also has the spinel structure, but has no divalent cations. It is the most commonly used material in the preparation of magnetic recording tapes. *See* CRYSTAL STRUCTURE; MAGNETIC RECORDING.

Properties. The important intrinsic parameters of a ferrite are the saturation magnetization, Curie temperature, and magnetocrystalline (K_1) and magnetostrictive (λ_s) anisotropies. These properties are determined by the choice of the cations and their distribution in the various sites. Commercially important ferrites invariably contain two or more M elements; and the iron content often deviates from stoichiometry to optimize the magnetic properties. High-permeability MnZn ferrites, for example, contain about 26% MnO, 22% ZnO, and 52% Fe_2O_3 ($Mn_{0.51}Zn_{0.44}Fe_{2.05}O_4$), at which composition both K_1 and λ_s are nearly zero.

In addition to the intrinsic magnetic parameters, microstructure plays an equally important role in determining device properties. Thus, grain size, porosity, chemical homogeneity, and foreign inclusions dictate in part such technical properties as permeability, line width, remanence, and coercivity in polycrystalline ceramics. In garnet films for bubble domain device applications, the film must essentially be free of all defects such as inclusions, growth pits, and dislocations.

Preparation. There are a number of methods by which ferrites may be prepared; the choice is dictated by the final form desired. Thin monocrystalline garnet films for bubble domain devices are grown by a liquid-phase epitaxy technique on top of a lattice- matched nonmagnetic gadolinium gallium garnet (GGG) crystal substrate. The GGG single crystals are usually grown by the Czochralski pulling technique. Spinel ferrite single crystals for use in recording heads are often grown by the Bridgman technique or the Czochralski technique. Oxide powders prepared for magnetic recording are obtained by precipitation from aqueous salt solutions. *See* SINGLE CRYSTAL.

Polycrystalline ferrites are most economically prepared by ceramic techniques. Component oxides or carbonates are mixed, calcined at elevated temperatures for partial compound formation, and then granulated by ball milling. Dispersants, plasticizers, and lubricants are added, and the resultant slurry is spray-dried, followed by pressing to desired shape and sintering. The last step completes the chemical reaction to the desired magnetic structure and effects homogenization, densification, and grain growth of the compact. It is perhaps the most critical step in optimizing the magnetic properties of commercial ferrites. *See* CERAMICS; FERRITE DEVICES. Gilbert Y. Chin

Bibliography. A. J. Baden-Fuller, *Ferrites at Microwave Frequencies*, 1987; D. J. Craik (ed.), *Magnetic Oxides*, 1975; J. Goldman, *Modern Ferrite Technology*, 1990; E. C. Snelling, *Soft Ferrites: Properties and Applications*, 2d ed., 1988; E. C. Snelling and A. D. Giles, *Ferrites for Inductors and Transformers*, 1984; C. Srivastava and M. J. Patni, *Ferrites*, 1990; R. Valenzuela, *Magnetic Ceramics*, 1994; H. Watanabe (ed.), *Ferrites*, 1982.

Ferrite devices

Electrical devices whose principle of operation is based upon the use and properties of ferrites, which are magnetic oxides. Ferrite devices are divided into two categories, depending on whether the ferrite is magnetically soft (low coercivity) or hard (high coercivity). Soft ferrites are used primarily as transformers, inductors, and recording heads, and in microwave devices. Since the electrical resistivity of soft ferrites is typically 10^6–10^{11} times that of metals, ferrite components have much lower eddy current losses and hence are used at frequencies generally above about 10 kHz. Hard ferrites are used in permanent-magnet motors, loudspeakers, and holding devices, and as storage media in magnetic recording devices. In this article, discussion of hard ferrites will be restricted to recording only. One type of soft ferrite, referred to as square-loop ferrites, was once deployed in huge quantities as cores for digital computer memories but has now been replaced by semiconductor integrated circuits. *See* FERRIMAGNETISM; FERRITE.

Chemistry and crystal structure. Modern soft ferrite devices stemmed from the contributions on spinel ferrites made by the Japanese and Dutch scientists during World War II. The general formula for the spinel is MFe_2O_4, in which M is a divalent metal ion. In special cases, the divalent ion M can be replaced by an equal molar mixture of univalent and trivalent ions. Thus lithium ferrite can be thought of as having the formula $Li_{0.5}Fe_{0.5}Fe_2O_4$. However, the commercially practical ferrites are those in which the divalent ion represents one or more magnesium (Mg), manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu), zinc (Zn), and cadmium (Cd) ions. The trivalent Fe ion may also be substituted by other trivalent ions such as aluminum (Al). The compositions are carefully adjusted to optimize the device requirements, such as permeability, loss, ferromagnetic resonance line width, and so forth.

The ferrimagnetic garnets were discovered in France and the United States in 1956. The general formula is $M_3Fe_5O_{12}$, in which M is a rare-earth or yttrium ion. Single-crystal garnet films form the basis of bubble domain device technology. Bulk garnets have applications in microwave devices. *See* FERRIMAGNETIC GARNETS.

Hard ferrites for permanent-magnetic device applications have the hexagonal magnetoplumbite structure, with the general formula $MFe_{12}O_{19}$, where M is usually barium (Ba) or strontium (Sr). Hexagonal ferrites exhibit large coercivity ($H_c \simeq 200$ kA/m or 2.5 kilooersteds) owing to a large magnetocrystalline anisotropy. Particulates of this material are also under study as potential storage media in magnetic recording.

The material of choice for magnetic recording is γ-Fe_2O_3, which has a spinel structure. The particles are usually modified with a thin cobalt layer on the surface to increase the coercivity for enhanced performance. *See* MAGNETIC RECORDING.

With the exception of some single crystals used in recording heads and special microwave applications, and particulates used as storage media in magnetic recording, all ferrites are prepared in polycrystalline form by ceramic techniques.

Applications. A summary of the applications of ferrites is given in the **table**. These may be divided into nonmicrowave, microwave, and magnetic recording applications. Further, the nonmicrowave applications may be divided into categories determined by the magnetic properties based on the *B-H* behavior, that is, the variation of the magnetic induction or flux density *B* with magnetic field strength *H*, as shown in the **illustration**. The plot in the illustration is termed a hysteresis loop, with the area encompassed by the loop being proportional to the power loss per unit volume within the ferrite. The categories are linear *B-H*, with low flux density, and nonlinear *B-H*, with medium to high flux density. The highly nonlinear *B-H*, with a square or rectangular hysteresis loop, was once exploited in computer memory cores.

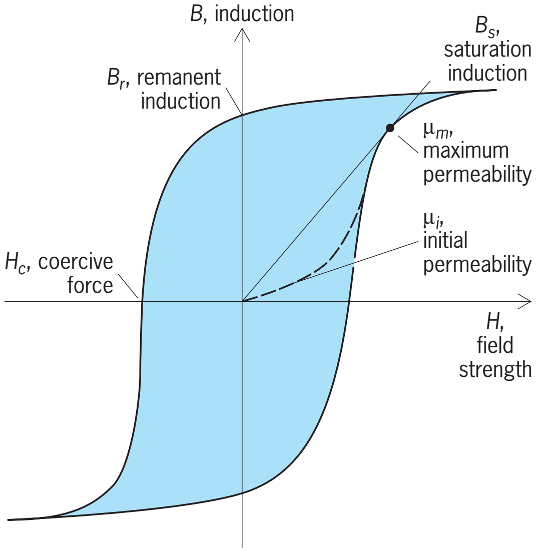
Linear B-H devices. In the linear region, the most important devices are high-quality inductors, particularly those used in filters in frequency-division multiplex telecommunications systems and low-power wide-band and pulse transformers. Virtually all such devices are made of either MnZn ferrite or NiZn ferrite, though predominantly the former.

In the design of inductors, the so-called μQ product of a material has been found to be a useful index of the quality of the material. In this product, μ is the initial permeability and Q is equal to $\omega L/R$,

Summary of ferrite applications

Ferrite chemistry	Device	Device function	Frequencies	Desired ferrite properties
		Linear B-H, low flux density[*]		
MnZn, NiZn	Inductor	Frequency selection network Filtering and resonant circuits	<1 MHz (MnZn) ~1–100 MHz (NiZn)	High μ, high μQ, high stability of μ with temperature and time
MnZn, NiZn	Transformer (pulse and wideband)	Voltage and current transformation Impedance matching	Up to 500 MHz	High μ low hysteresis losses
NiZn	Antenna rod	Electromagnetic wave receival	Up to 15 MHz	High μQ, high resistivity
MnZn	Loading coil	Impedance loading	Audio	High μ high B_s, high stability of μ with temperature, time, and dc bias
		Nonlinear B-H, medium-to-high flux density[*]		
MnZn, NiZn	Flyback transformer	Power converter	<100 kHz	High μ, high B_s, low hysteresis losses
MnZn	Deflection yoke	Electron-beam deflection	<100 kHz	High μ, high B_s
MnZn, NiZn	Suppression bead	Block unwanted ac signals	Up to 250 MHz	Moderately high μ, high B_s, high hysteresis losses
MnZn, NiZn	Choke coil	Separate ac from dc signals	Up to 250 MHz	Moderately high μ, high B_s, high hysteresis losses
MnZn, NiZn	Recording head	Information recording	Up to 10 MHz	High μ, high density, high μQ, high wear resistance
MnZn	Power transformer	Power converter	up to 200 kHz	High B_s, low hysteresis losses
		Microwave properties		
YIG[†], MgMn[‡], Li[§], NiZn	Isolators, attenuators, circulators, switches, modulators	Impedance matching, power level control, power splitting	1–5 GHz (YIG), 2–30 GHz (YIG, MgMn, Li), 30–100 GHz (NiZn)	Controlled B_s, high resistivity, high Curie temperature, narrow resonance line–width
		Magnetic recording properties		
γ-Fe_2O_3, Co-γ-Fe_2O_3	Audio and video tapes, computer disks	Information storage		High B_s, high H_c

[*] After P. I. Slick, Ferrites for nonmicrowave applications, in E. P. Wohlfarth (ed.), *Ferromagnetic Materials*, vol. 2, North-Holland, 1980.
[†] May contain Al, Gd.
[‡] May contain Al, Zn.
[§] May contain Ti, Zn.



Plot of magnetic induction *B* as a function of magnetic field strength *H*.

where ω is the angular frequency, L the inductance, and R the effective series resistance arising from core loss. The higher the value of ωQ, the better is the material, and ferrites have the highest ωQ product of any commercially available magnetic material. Typical ωQ values measured at 20 kHz for various materials are iron dust, 2000; Permalloy powder, 10,000; 12.5-micrometer-thick molybdenum Permalloy tape, 100,000; and MnZn ferrite, 250,000. Values of ωQ greater than 10^6 at 20 kHz have been achieved commercially in specially prepared MnZn ferrites.

As compared with NiZn ferrites, MnZn ferrites have lower residual and hysteresis losses, higher permeabilities, and lower resistivities. A high permeability lowers the frequency of ferromagnetic resonance, which is accompanied by a large rise in losses. This factor, along with increased eddy current losses coming from lower resistivities, restricts the useful upper frequency range for MnZn ferrites in high-Q inductors to about 1 MHz. Above that, NiZn ferrites are preferred. *See* ELECTRIC FILTER; INDUCTOR; Q (ELECTRICITY); TRANSFORMER.

For transformer applications the highest value of ω over the operating frequency range is desired. Values of ω in the 18,000 range (10 kHz) for MnZn ferrite are commercially available, although values up to 40,000 have been obtained in the laboratory.

Nonlinear B-H devices. The largest usage of ferrite measured in terms of material weight is in the nonlinear *B-H* range, and is found in the form of deflecting yokes and flyback transformers for television receivers. The cores for these devices must have high saturation induction B_s along with high maximum permeability ω_m at the knee of the *B-H* curve to frequencies as high as 100 kHz, the effective flyback frequency used in scanning a television tube. Again, MnZn and NiZn ferrites dominate the use in these devices. *See* TELEVISION RECEIVER.

A rapidly growing use of ferrites is in the power area, where ferrite transformers are extensively used in switched mode (alternating current to direct current) and converter mode (direct current to direct current) power supplies. Such power supplies are widely used in various computer peripheral equipment and private exchange telephone systems. Here a large value of B_s together with low hysteresis losses is important. MnZn ferrites are superior to NiZn ferrites in both aspects. *See* ELECTRONIC POWER SUPPLY.

Although growth area for ferrites in the nonlinear *B-H* range is in recording heads. As compared with metals, such as Permalloy and Sendust, ferrites have higher electrical resistivity and wear resistance and are hence preferred in video and high-frequency recording applications.

Microwave devices. Microwave devices make use of the nonreciprocal propagation characteristics of ferrites close to or at a gyromagnetic resonance frequency in the range of 1–100 GHz. The most important of such devices are isolators and circulators. Materials having a range of B_s are needed for operation at various frequencies since for resonance, B_s is less than ω/γ where γ is the gyromagnetic ratio. In the 1–5-GHz range, yttrium iron garnet of $B_s = 0.02$–0.18 tesla is used. In the 2–30-GHz range, MgMn, MgMnZn, MgMnAl, and Li ferrites of $B_s = 0.06$–0.25 T are used along with the garnets. At 30–100 GHz, NiZn ferrites with B_s up to 0.50 T are used. The garnets have highly desirable, small, ferromagnetic-resonance linewidths, particularly in single-crystal form. In device development there is a strong trend toward realizing the conventional waveguide components in microstrip form whereby both the transmission and the gyromagnetic function are provided by a ferrimagnetic substrate, or the gyromagnetic function is provided by a ferrimagnetic insert on a nonmagnetic ferrite substrate. *See* GYRATOR; MICROWAVE.

Magnetic recording devices. Vast amounts of audio and video information and digital data from computers are stored in magnetic tapes and disks. Here magnetic recording materials function as hard magnetic materials with coercivity in the 25–100 kA/m (310–1250 oersteds) range. The materials are generally prepared in particulate form by chemical precipitation techniques and then dispersed in an organic slurry, either on a plastic substrate (tape or flexible disk) or on a rigid disk (usually an Al-Mg alloy). In a more advanced technique, continuous films are deposited by using sputtering or evaporation, particularly metal alloys rather than ferrites, on the smaller rigid disks. The most widely used particles in magnetic recording are γ-Fe_2O_3 and co-modified γ-Fe_2O_3. The basic γ-Fe_2O_3 is generally used in audio recording, while the higher-coercivity Co–γ-Fe_2O_3 [$H_c \simeq$ 50 kA/m (620 Oe) versus 25 kA/m (310 Oe) for γ-Fe_2O_3] dominates video recording. The hexagonal ferrites, Ba(or Sr)$Fe_{12}O_{19}$, are under study for use in magnetic recording. Since the easy axis of magnetization is normal to the plane of the particle platelets, such particles may be used as perpendicular recording media, in contrast to the normal longitudinal mode. *See* COMPUTER STORAGE TECHNOLOGY.

Gilbert Y. Chin

Bibliography. J. Goldman, *Modern Ferrite Technology*, 1990; National Advisory Board, *Report of the Committee on Magnetic Materials*, Publ. NMAB-426, National Research Council, 1985; E. C. Snelling, *Soft Ferrites: Properties and Applications*, 2d ed., 1988; E. C. Snelling and A. D. Giles, *Ferrites for Inductors and Transformers,* 1983; G. Winkler, *Magnetic Garnets*, 1981; E. P. Wohlfarth (ed.), *Ferromagnetic Materials*, vol. 2, 1980.

Ferroalloy

An important group of metallic raw materials required for the steel industry. Ferroalloys are the principal source of such additions as silicon, Si, and manganese, Mn, which are required for even the simplest plain-carbon steels; and chromium, Cr, vanadium, V, tungsten, W, titanium, Ti, and molybdenum, Mo, which are used in both low- and high-alloy steels. Also included are many other more complex alloys. Ferroalloys are unique in that they are brittle and otherwise unsuited for any service application, but they are important as the most economical source of these elements for use in the manufacture of the engineering alloys. These same elements can also be obtained, at much greater cost in most cases, as essentially pure metals. The ferroalloys contain significant amounts of iron and usually have a lower melting range than the pure metals and are therefore dissolved by the molten steel more readily than the pure metal. In other cases, the other elements in the ferroalloy serve to protect the critical element against oxidation during solution and thereby give higher recoveries. Ferroalloys are used both as deoxidizers and as a specified addition to give particular properties to the steel. *See* STEEL MANUFACTURE.

Many ferroalloys contain combinations of two or more desirable alloy additions, and well over 100 commercial grades and combinations are available. Although of less general importance, other sources of these elements for steelmaking are metallic nickel, Ni, silicon carbide, molybdic oxide, and even misch metal (a mixture of rare earths). Analyses of a few typical ferroalloys are given in the **table**.

The three ferroalloys which account for the major tonnage in this class are the various grades of silicon, manganese, and chromium. For example, 13 lb (5.9 kg) of manganese is used on the average in the United States for every ton of open-hearth steel produced. Elements supplied as ferroalloys are among the most difficult metals to reduce from ore.

The most common grade of ferromanganese is a blast-furnace product (standard ferromanganese), the major variation from pig iron production being the use of manganese-rich ore. Other grades of ferromanganese with low C are made by the processes described in the next sections. A low-Si ferroalloy can be made in the same way, but the tonnage grades are made by other methods.

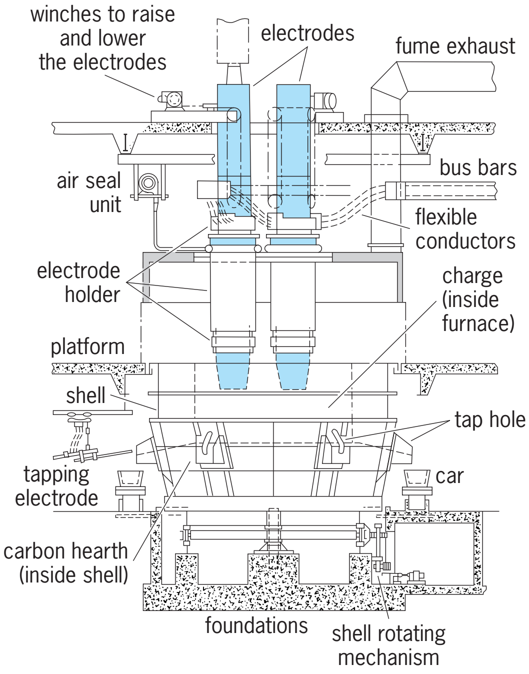
The most general method of ferroalloy manufacture is the submerged-arc furnace (see **illus.**). Its use will be described in connection with the production of 50% ferrosilicon. A modern furnace of this type is 26 ft (7.9 m) in diameter and 10 ft (3 m) deep with the three carbon electrodes supplying an average of 5 kWh per pound (2.4×10^8 joules per kilogram) of silicon produced. The furnace produces about 2 tons (1.8 metric tons) of alloy per hour. The furnace is charged intermittently from the top with a mixture of the required amounts of SiO_2 (quartzite rock), C (coke), and Fe (as turnings), and the ferrosilicon is periodically tapped from the bottom. The overall operation is essentially continuous, like that of a blast furnace, but the heat is supplied by the resistance of the charge and some arcing, and the coke serves primarily as a reducing agent. After solidification in flat molds, the ferroalloy is crushed to specified sizes for delivery. The overall reaction for the process can be

Analysis of typical ferroalloys, % weight

Type of ferroalloy*	Mn	Si	C	Cr	Mo	Al	Ti	V
Ferromanganese								
Standard	78–82	1.25	7.5					
Medium carbon	80–85	1.25–2.5	1–3†					
Low carbon	80–85	1.25–7.0	0.75					
Ferrosilicon								
50% regular	—	47–52	0.15					
75% regular	—	73–78	0.15					
Ferrochromium								
High carbon	—	1–2	4.5–6.0†	67–70				
Low carbon	—	0.3–1.0	0.03–2.0†	68–71				
SM low carbon	4–6	4–6	1.25	62–65				
Ferromolybdenum								
High carbon	—	1.5	2.5	—	55–70			
Ferrovanadium								
High carbon	—	13.0	3.5	—	—	1.5	—	30–40
Ferrotitanium								
Low carbon	—	3–5	0.1	—	—	6–10	38–43	

[*] In all cases the balance is Fe, with the exception of minor impurities. The latter are usually specified, such as 0.10% max P.

[†] In several specified grades within this range.



Submerged-arc furnace used in ferroalloy manufacture. (*Lectromelt Furnace Division*, *McGraw-Edison Co.*)

written as (1). The Si reduced in this way dissolves in

$$SiO_2 + C \rightarrow Si + 2CO \qquad (1)$$

the molten iron which is also present and simplifies the operation of the process.

Because SiO_2 is present as gangue in many ores, and because silicate slags have sufficiently low melting points to be controlled conveniently, many other ferroalloys contain appreciable amounts of Si when they are manufactured by this same process. The inherent low solubility of C in Si alloys makes this combination desirable when Si is not objectionable. Low-silicon grades can be made, however, by further addition of metal oxide as illustrated by the general reaction (2). Thus ferrosilicon can be used as an

$$2MO + Si \rightarrow SiO_2 + 2M \qquad (2)$$

intermediate to produce other ferroalloys, and this scheme is common.

Likewise, aluminum can be used as the reducing agent by a process called the thermit reaction, as indicated in (3), in which all or part of the heat required

$$2Al + Fe_2O_3 \rightarrow 2Fe + Al_2O_3 + \text{heat} \qquad (3)$$

results from the reduction reaction. The oxide of various desirable alloys may be substituted for Fe_2O_3, and Ca or Mg may be used for Al, a variety of such combinations being possible. The product is low in both Si and C.

Thus, depending upon the reducibility of the ore, and the amounts of C, Si, and Fe that can be tolerated in the product, as well as economic considerations, ferroalloys are produced in the blast furnace, or in the submerged-arc or similar variations of the electric furnace, or by aluminothermic or silicothermic reactions. These ferroalloys are then supplied to the steel industry for use as deoxidizers or alloy-addition agents. *See* ARC HEATING; IRON ALLOYS; MANGANESE; MOLYBDENUM; SILICON; STEEL; VANADIUM. Gerhard Derge

Bibliography. P. D. Deeley et al., *Ferroalloys and Alloying Handbook*, 1981; R. Serjeantson, *Ferro-Alloy Directory and Databook*, 3d ed., 1992; Y. S. Touloukian and C. Y. Ho, *Properties of Selected Ferrous Alloying Elements*, vols. 1–3, 1989.

Ferroelectrics

Crystalline substances which have a permanent spontaneous electric polarization (electric dipole moment per cubic centimeter) that can be reversed by an electric field. In a sense, ferroelectrics are the electrical analog of the ferromagnets, hence the name. The spontaneous polarization is the so-called order parameter of the ferroelectric state, just as the spontaneous magnetization is the order parameter of the ferromagnetic state. The names Seignette-electrics or Rochelle-electrics, which are also widely used, are derived from the name of the first substance found to have this property, Seignette salt or Rochelle salt. *See* FERROMAGNETISM.

The reversibility of the spontaneous polarization is due to the fact that the structure of a ferroelectric crystal can be derived from a nonpolarized structure by small displacements of ions. In most ferroelectric crystals, this nonpolarized structure becomes stable if the crystal is heated above a critical temperature, the ferroelectric Curie temperature; that is, the crystal undergoes a phase transition from the polarized phase (ferroelectric phase) into an unpolarized phase (paraelectric phase). The change of the spontaneous polarization at the Curie temperature can be continuous or discontinuous. The Curie temperature of different types of ferroelectric crystals range from a few degrees absolute to a few hundred degrees absolute. As a rule, the ferroelectric phase is the low-temperature phase; however, there are crystals which are ferroelectric in a relatively narrow temperature range only, and others stay polarized up to the temperature of decomposition or melting.

Classification. From a practical standpoint ferroelectrics can be divided into two classes. In ferroelectrics of the first class, spontaneous polarization can occur only along one crystal axis; that is, the ferroelectric axis is already a unique axis when the material is in the paraelectric phase. Typical representatives of this class are Rochelle salt, KH_2PO_4, $(NH_4)_2SO_4$, guanidine aluminum sulfate hexahydrate, glycine sulfate, colemanite, and thiourea.

In ferroelectrics of the second class, spontaneous polarization can occur along several axes that are equivalent in the paraelectric phase. The following substances, which are all cubic above the Curie point, belong to this class: $BaTiO_3$-type (or perovskite-type) ferroelectrics; $Cd_2Nb_2O_7$; $PbNb_2O_6$;

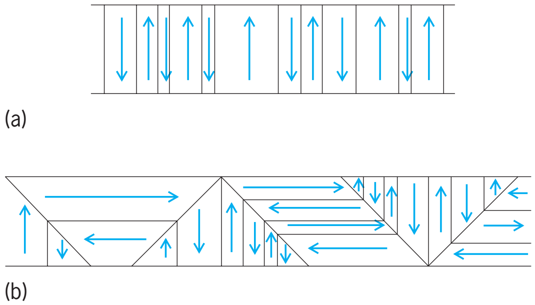


Fig. 1. Domain configurations (simplified) encountered in ferroelectric crystals. (*a*) First class. (*b*) Second class.

certain alums, such as methyl ammonium alum; and $(NH_4)_2Cd_3(SO_4)_3$. Some of the $BaTiO_3$-type ferroelectrics have, below the Curie temperature, additional transition temperatures at which the spontaneous polarization switches from one crystal axis to another crystal axis. For example, $BaTiO_3$ and $KNbO_3$ polarize with decreasing temperature first along a [100] axis, then the polarization switches into a [110] axis, and finally into a [111] axis.

From a scientific standpoint, one can distinguish proper ferroelectrics and improper ferroelectrics. In proper ferroelectrics, for example, $BaTiO_3$, KH_2PO_4, and Rochelle salt, the spontaneous polarization is the order parameter. The structure change at the Curie temperature can be considered a consequence of the spontaneous polarization. The unit cell of the crystal in the ferroelectric phase contains the same number of chemical formula units as the unit cell in the paraelectric phase. In improper ferroelectrics, the spontaneous polarization can be considered a by-product of another structural phase transition. The unit cell in the ferroelectric phase is an integer multiple of the unit cell in the paraelectric phase. Examples of such systems are $Gd(MoO_4)$ and boracites. The dielectric elastic and electromechanical behavior of the two types of ferroelectrics differ significantly.

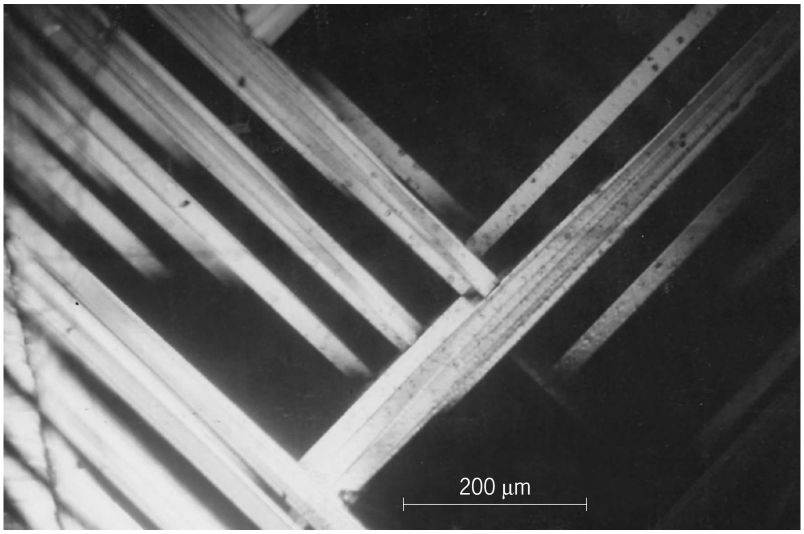


Fig. 2. Ferroelectric domains in $BaTiO_3$ photographed through a polarizing microscope. Ferroelectric domains range from macroscopic to submicroscopic size.

Ferroelectric domains. The spontaneous polarization can occur in at least two equivalent crystal directions; thus, a ferroelectric crystal consists in general of regions of homogeneous polarization that differ only in the direction of polarization. These regions are called ferroelectric domains. Ferroelectrics of the first class consist of domains with parallel and antiparallel polarization (**Fig. 1***a*), whereas ferroelectrics of the second class can assume much more complicated domain configurations (Fig. 1*b*). The region between two adjacent domains is called a domain wall. Within this wall, the spontaneous polarization changes its direction. The wall between antiparallel domains is probably only a few lattice spacings thick, whereas the wall between domains polarized at a right angle to each other is probably thicker. Ferroelectric domains can be observed in a number of substances by means of the polarizing microscope (**Fig. 2**) because of their birefringence, or double refraction. The ferroelectric domains range in size from macroscopic (millimeters) to submicroscopic. *See* BIREFRINGENCE.

Ferroelectric hysteresis. When an electric field is applied to ferroelectric crystal, the domains that are favorably oriented with respect to this field grow at the expense of the others, for example, by sidewise motion of domain walls. In addition, favorably oriented domains can nucleate and grow until the whole crystal becomes one single domain. When the field is reversed, the polarization reverses through the same processes. The relation between the resulting polarization P of the whole crystal and the externally applied electric field E is given by a hysteresis loop (**Fig. 3**). The shape of the hysteresis loop depends strongly upon the perfection of the crystal as well as upon the rate of change of the externally applied field E. A simple circuit that permits observation of ferroelectric hysteresis loops with an oscilloscope is shown in **Fig. 4**. In some ferroelectrics, polarization can be reversed within a fraction of 1 microsecond.

Spontaneous polarization. The magnitude of the permanent or spontaneous polarization P_s of a domain can be obtained from the hysteresis loop by extrapolating the saturation branch to zero external field (Fig. 3). For most ferroelectrics, the values of P_s are between 10^{-7} and 10^{-4} coulomb/cm^2 (**Fig. 5**). In nonferroelectric dielectrics, electric fields between 10^5 and 10^8 V/cm would be necessary in order to achieve such large polarizations.

Dielectric properties. As a rule, the dielectric constant ϵ measured along a ferroelectric axis increases in the paraelectric phase when the Curie temperature is approached. In many ferroelectrics, this increase can be approximated by the Curie-Weiss law, shown in the equation below. Here T designates the

$$\epsilon = \frac{C}{T - T_0}$$

temperature of the crystal, and T_0 is equal to or somewhat smaller than the transition temperature. C is the so-called Curie constant. For $BaTiO_3$, this law holds

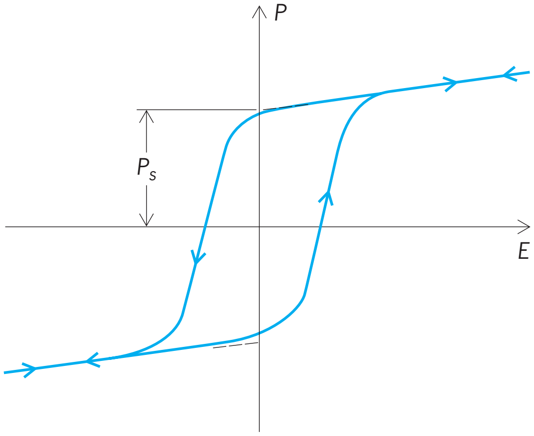


Fig. 3. Net polarization *P* of a ferroelectric crystal versus externally applied electric field *E*.

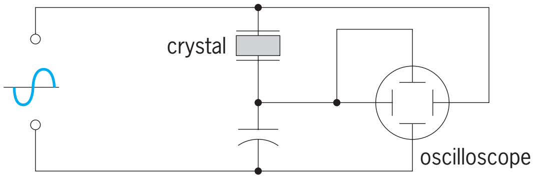


Fig. 4. Circuit for the display of ferroelectric hysteresis loops on an oscilloscope.

unaltered up to frequencies of 2.4×10^{10} Hz. Dispersion sets in the far-infrared. The dielectric constant drops when the crystal becomes spontaneously polarized (**Fig. 6**). In the ferroelectric phase, the dielectric constant has two components. The first component is the dielectric constant of the individual domains. It is independent of the frequency and of the electric field generally up to far-infrared frequencies. The second component is due to domain wall motions, that is, to partial reversal of the spontaneous polarization. This process can give rise to large dielectric losses, and it depends strongly upon the frequency, the electric field strength, the domain structure, and the temperature. In uniaxial ferroelectrics, the dielectric constant measured perpendicular to the ferroelectric axis generally does not show a very pronounced anomaly near the Curie temperature, and in some cases it has even the same order of magnitude and temperature dependence as for any normally behaving dielectric crystal. *See* CURIE-WEISS LAW; DIELECTRIC MATERIALS; PERMITTIVITY.

Piezoelectric properties. Ferroelectrics can be divided into two groups according to their piezoelectric behavior.

The ferroelectrics in the first group are already piezoelectric in the unpolarized phase. Those piezoelectric moduli which relate stresses to polarization along the ferroelectric axis have essentially the same temperature dependence as the dielectric constant along this axis, and hence become very large near the Curie point. The spontaneous polarization gives rise to a large spontaneous piezoelectric strain which is proportional to the spontaneous polarization. In KH_2PO_4-type ferroelectrics and in Rochelle salt, for example, this strain is a shear in the plane perpendicular to the axis of polarization. It reaches 27 min of arc in KH_2PO_4 and about 1.8 min of arc in Rochelle salt. The piezoelectric modulus decreases as the spontaneous polarization increases. But with sufficiently large stresses, it is possible to align the domains and reverse the spontaneous polarization (**Fig. 7**). The relation between the resulting polarization of the whole crystal and the mechanical stress is given by a hysteresis loop analogous to the loop of Fig. 3 (piezoelectric hysteresis). This effect can simulate a very large piezoelectric modulus.

The ferroelectrics in the second group are not piezoelectric when they are in the paraelectric phase. However, the spontaneous polarization lowers the symmetry so that they become piezoelectric in the polarized phase. This piezoelectric activity is often hidden because the piezoelectric effects of the

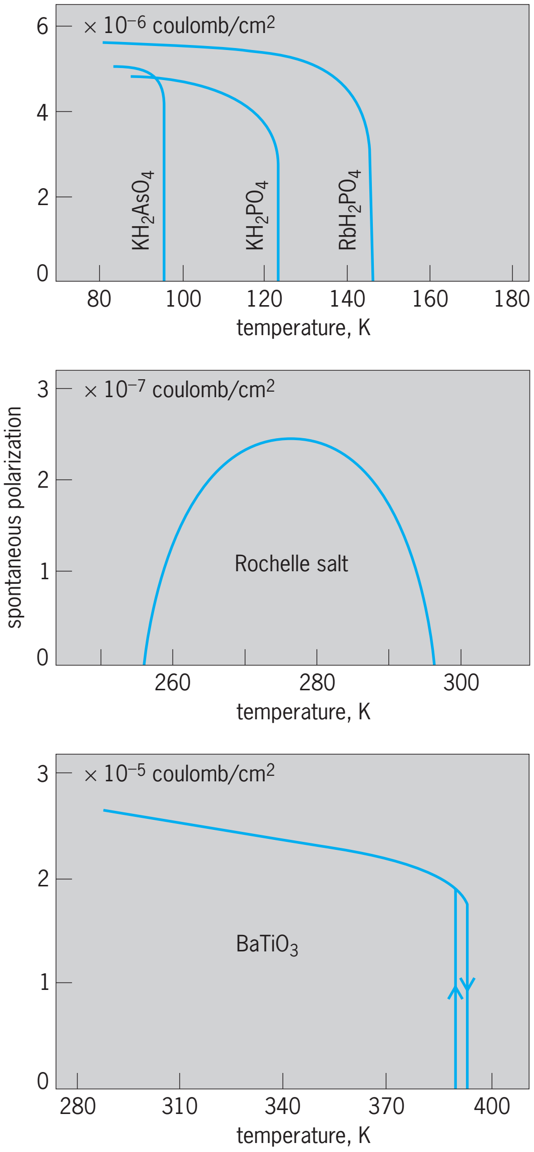


Fig. 5. Dependence upon temperature of the spontaneous polarization of some ferroelectrics. °F = (K × 1.8) − 459.67. °C = K − 273.15.

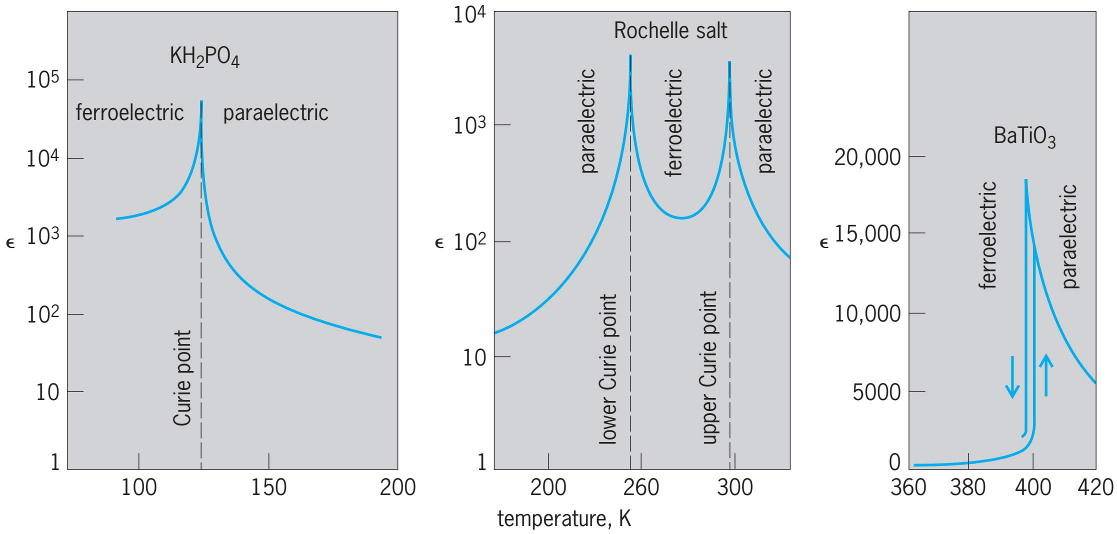


Fig. 6. Anomalous temperature dependence of relative dielectric constant of ferroelectrics at transition temperature. °F = (K × 1.8) − 459.67. °C = K − 273.15.

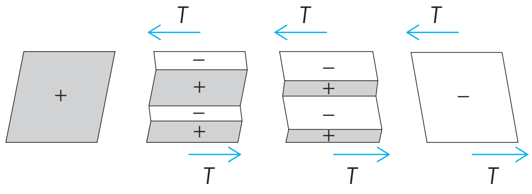


Fig. 7. Schematic representation of the reversal of the spontaneous polarization by a mechanical shear stress *T* in KH_2PO_4 and Rochelle salt.

various domains can cancel. However, strong piezoelectric activity of a macroscopic crystal or even of a polycrystalline sample occurs when the domains have been aligned by an electric field. The spontaneous strain is proportional to the square of the spontaneous polarization. In $BaTiO_3$, for example, the crystal (which has cubic symmetry in the unpolarized phase) expands along the axis of polarization and contracts at right angles to it. The strain is of the order of magnitude of 1%. The spontaneous polarization cannot be reversed by a mechanical stress in ferroelectrics of this group. *See* PIEZOELECTRICITY.

Crystal structure. The structures of different types of ferroelectrics are entirely different, and it is not possible to establish a general rule for the occurrence of ferroelectricity. The structures of a number of ferroelectrics and the minute changes that they undergo when spontaneous polarization occurs are known in great detail from x-ray diffraction and neutron diffraction studies. In a qualitative way, the process of polarization is best understood for ferroelectrics of the $BaTiO_3$ type. **Figure 8** shows schematically the structure of the unit cell of a $BaTiO_3$ crystal in the unpolarized state, and the arrows indicate the direction in which the ions are slightly displaced when the lattice becomes spontaneously polarized along the axis *z*. The order of magnitude of the displacement is 1% of the unit cell dimension. However, these displacements do not account quantitatively for the observed polarization, because other changes of the electronic structure occur as well. *See* CRYSTAL STRUCTURE.

In KH_2PO_4-type ferroelectrics, hydrogen bonds O—H · · · O play an important part in the ferroelectric effect. Above the Curie temperature, the hydrogen ions are statistically distributed over the two possibilities O—H · · · O and O · · · H—O, whereas below the Curie point, one or the other of these two possibilities is strongly favored, depending upon the sign of the spontaneous polarization.

Antiferroelectric crystals. These materials are characterized by a phase transition from a state of lower symmetry (generally low-temperature phase) to a state of higher symmetry (generally high-temperature phase). The low-symmetry state can be regarded as a slightly distorted high-symmetry state. It has no permanent electric polarization, in contrast to ferroelectric crystals. The crystal lattice can be

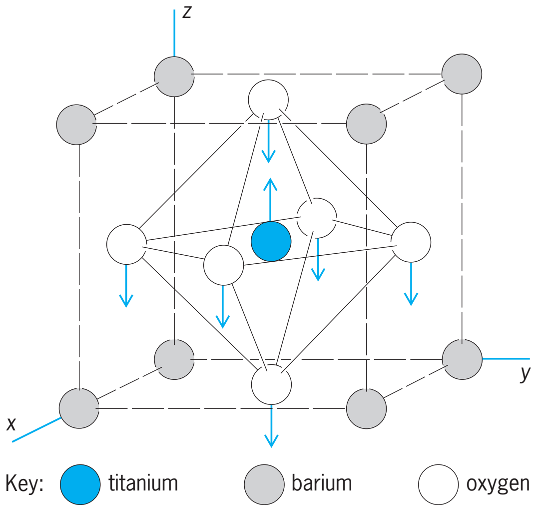


Fig. 8. Crystal lattice of $BaTiO_3$. Arrows indicate displacements of the ions when crystal becomes polarized.

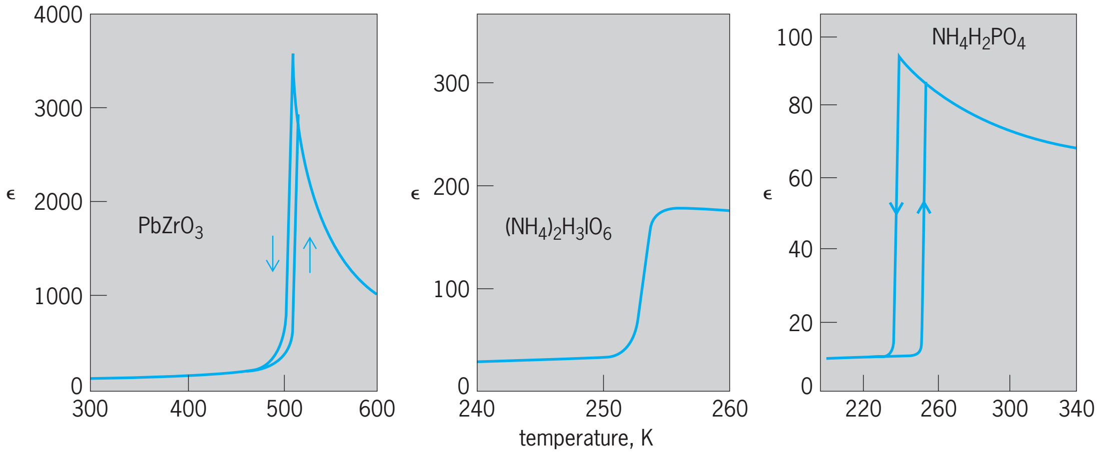


Fig. 9. Anomalous temperature dependence of relative dielectric constant of antiferroelectric crystals. Note increase as transition temperature is approached and drop when antipolarization occurs. °F = (K × 1.8) − 459.67. °C = K − 273.15.

regarded as consisting of two interpenetrating sublattices with equal but opposite electric polarization. This state is referred to as the antipolarized state.

In a certain sense, an antiferroelectric crystal is the electrical analog of an antiferromagnetic crystal. In the high-symmetry phase, the sublattices are unpolarized and indistinguishable. In general, antiferroelectric crystals have more than one axis along which the sublattices can polarize. Therefore, the low-symmetry phase consists of regions of homogeneous antipolarization which differ only in the orientation of the axis along which antipolarization has occurred. These regions are called antiferroelectric domains and can be observed by the polarizing microscope. Because these domains have no permanent electric dipole moment, an electric field generally has little influence on domain structure. *See* ANTIFERROMAGNETISM.

The dielectric constant of antiferroelectric crystals is generally larger than it is for nonferroelectric crystals and has an anomalous temperature dependence. It increases as the transition temperature is approached and drops when antipolarization occurs (**Fig. 9**). In some antiferroelectrics the phase transition is discontinuous; in others it is continuous.

The structure of antiferroelectric crystals is generally closely related to the structure of ferroelectric crystals. Some antiferroelectrics even undergo phase transitions from an antipolarized state into a spontaneously polarized, ferroelectric state; in others a sufficiently strong electric field applied along an antiferroelectric axis reverses the polarity of one of the sublattices so that a ferroelectric state results. The crystal reverts, however, to the antiferroelectric state when the electric field is removed. **Figure 10** shows net polarization versus externally applied field for such a case.

Compounds with antiferroelectric properties are $PbZrO_3$, $PbHfO_3$, $NaNbO_3$ (isomorphous with ferroelectric $BaTiO_3$), WO_3 (structure related to $BaTiO_3$), $NH_4H_2PO_4$ and isomorphous NH_4 salts (isomorphous with ferroelectric KH_2PO_4), $(NH_4)_2H_3IO_6$, $Ag_2H_3IO_6$, and certain alums.

Origin of phase transition. The ferroelectric phase transition results from an instability of one of the normal lattice vibration modes. On approaching the transition temperature, the frequency of the relevant normal mode decreases (soft mode). The restoring force of the mode displacements tends to zero. When the stability limit is reached, the displacements corresponding to the soft mode freeze in, and the ferroelectric phase results. The ferroelectric soft mode is polar (infrared-active) and of infinite wavelength. The antiferroelectric phase transition, on the other hand, emerges from a soft lattice mode with a finite wavelength equal to an integer multiple of a lattice period.

Applications. The piezoelectric effect of ferroelectrics (and certain antiferroelectrics) finds numerous applications in electromechanical transducers. The large electrooptical effect (birefringence

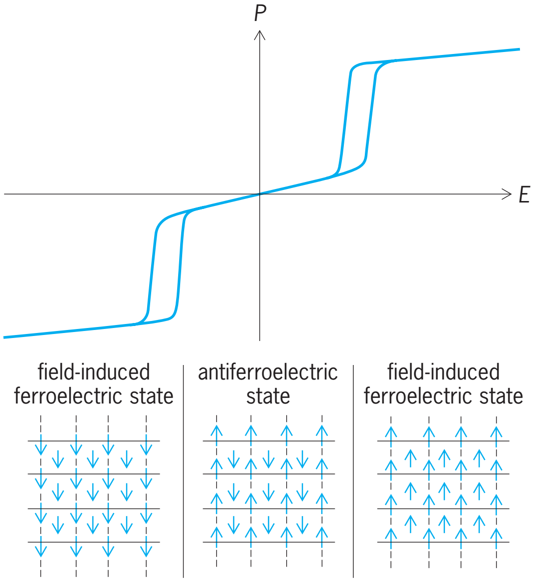


Fig. 10. Polarization *P* of antiferroelectric $PbZrO_3$ versus externally applied electric field *E*. Strong fields "switch" the antiferroelectric crystal into a ferroelectric state, as shown here schematically.

induced by an electric field) is used in light modulators. In some ferroelectrics (such as $BaTiO_3$, $LiNbO_3$, $KTaNbO_3$, and $LiTaO_3$), light can induce changes of the refractive indices. These substances are used for optical information storage and in real-time optical processors. The temperature dependence of the spontaneous polarization corresponds to a strong pyroelectric effect that can be exploited in thermal and infrared sensors. Werner Kanzig

Bibliography. J. M. Herbert, *Ferroelectric Transducers and Sensors*, 1982; F. Jona and G. Shirane, *Ferroelectric Crystals*, 1962, reprint 1993; M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials*, 1977; T. Mitsui, *An Introduction to the Physics of Ferroelectrics*, 1976; A. Moulson and J. M. Herbert, *Electroceramics: Materials, Properties, Applications*, 1992; N. Setter and E. L. Colla (eds.), *Ferroelectric Ceramics*, 1992; G. A. Smolenskii, *Ferroelectrics and Related Materials*, 1984; X. Yuhuan, *Ferroelectric Materials and Their Applications*, 1991.

Ferrofluid

A type of magnetic fluid that consists of a colloidal suspension of nanoscopic magnetic particles in a carrier liquid. A ferrofluid behaves as a fluid that is attracted to an external magnetic field without a large change in its viscosity.

Magnetic fluids are liquids that respond in many interesting ways to external magnetic fields. Most magnetic fluids combine conventional liquids (oil or water) with very small magnetic particles. Magnetorheological fluids are oils loaded with micrometer-sized magnetic particles. These materials experience a large change in viscosity, going from a thick fluid to a peanut butter consistency, when exposed to a magnetic field. Applications include computer-controlled vehicle suspension systems, where the damping in shock absorbers is controlled by magnetically regulating the magnetorheological fluid viscosity inside the shock absorber. Ferrofluids are similar to magnetorheological fluids except the particle sizes are 2–3 orders of magnitude smaller. These magnetic particles are so small (~10 nanometers in diameter) that thermal agitation in the fluid is great enough to keep the particles in suspension. One significant difference between magnetorheological fluids and ferrofluids is that ferrofluids behave as a colloidal suspension. The particles are so small and finely divided that they will not settle under the action of gravity or magnetic fields. Unlike magnetorheological fluids, ferrofluids experience a much smaller change in viscosity when exposed to magnetic fields. *See* COLLOID; MAGNETIC FIELD.

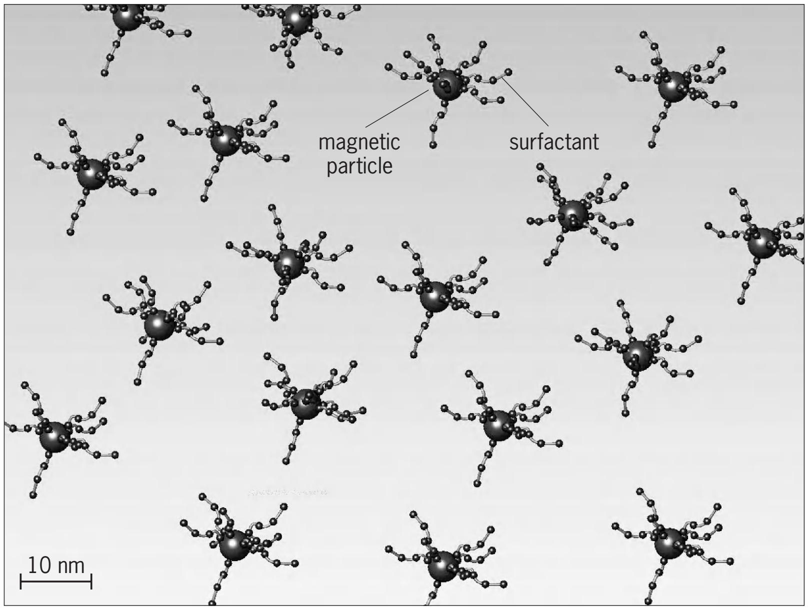


Fig. 1. Basic composition of a ferrofluid—magnetic nanoparticles coated with surfactant in a carrier fluid.

Composition. Ferrofluids consist of three basic components: a carrier fluid, magnetic nanoparticles, and a surfactant covering the nanoparticle (**Fig. 1**). Carrier fluids are generally organic solvents or water. Oil-based ferrofluids are used in applications such as magnetic seals, heat conduction for loudspeakers, and damping in stepper motors. Water-based ferrofluids are finding many useful biomedical applications such as magnetic cell sorting and biochemical sensing. The most common type of magnetic nanoparticle is magnetite (Fe_3O_4). The particle size is so small that the particle is considered superparamagnetic, meaning that there is only one magnetic domain in the particle. When an external field is applied to the fluid, the particles line up along the same field lines (**Fig. 2**). When the field is removed, thermal agitation in the fluid is great enough to cause the particles to resume their random orientation. The primary role of the surfactant is to aid in dispersing the particles within the carrier fluid. One end of the surfactant is absorbed on the particle surface, while the other interacts with the fluid and other particles. *See* MAGNETITE; NANOPARTICLES; SOLVENT; SURFACTANT.

There are a number of different methods for creating magnetite nanoparticles. The most common methods are ball grinding and chemical coprecipitation. In ball grinding, small particles are made from larger ones; whereas in chemical coprecipitation, nanoparticles are made in solution. Ball grinding, or size reduction, is a wet grinding process in which larger magnetite particles are mixed with a surfactant and rubbed together for long periods (~1000 hours). Chemical coprecipitation is based on mixing solutions of divalent and trivalent iron salts. As sodium hydroxide is dripped into the solution, magnetite nanoparticles precipitate out. This process requires careful control of the solution concentration, temperature, and pH. *See* PRECIPITATION (CHEMISTRY).

Applications. Most applications of ferrofluids exploit the combined fluid and magnetic behavior of the fluid. Most electromagnetic devices (motors, speakers, generators) are thermally limited. Some high-quality speakers use ferrofluids to improve

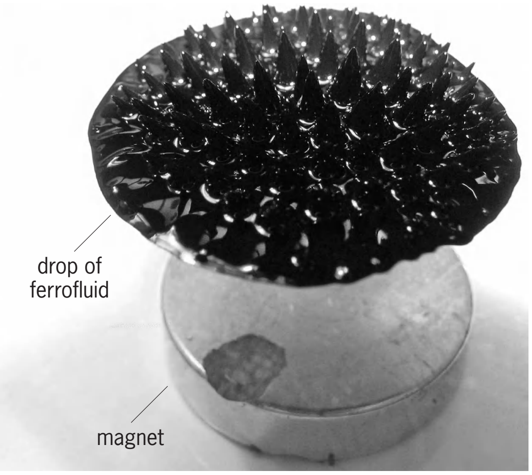


Fig. 2. Ferrofluid drop exposed to a magnetic field showing spikes due to magnetic field lines.

performance. The ferrofluid is captured magnetically in the air gap of the voice coil. The fluid serves two purposes: increased thermal conductivity to provide greater power capacity (reduce magnet and coil size) and mechanical damping to reduce harmonic distortion. Many stepper motors and optical disc actuators for CD and DVD players experience high-frequency vibration or ringing. A small volume of ferrofluid in the air gap of these actuators provides a very simple means of damping high-frequency vibration without additional design features. Ferrofluid seals, based on magnetically containing a small volume of ferrofluid between two concentric magnetic components, provides a liquid O-ring that serves as a hermetic seal with no friction and long life. *See* LOUDSPEAKER; STEPPING MOTOR. Lonnie J. Love

Bibliography. B. Berkovskii and V. Bashtovoi, *Magnetic Fluids and Applications Handbook*, Begell House, 1996; E. Blums, A. Cerbers, and M. Maiorov, *Magnetic Fluids*, Walter de Gruyter, 1997; R. E. Rosensweig, *Ferrohydrodynamics*, Dover, 1985.

Ferromagnetism

A property exhibited by certain metals, alloys, and compounds of the transition (iron group), rare-earth, and actinide elements in which, below a certain temperature called the Curie temperature, the atomic magnetic moments tend to line up in a common direction. Ferromagnetism is characterized by the strong attraction of one magnetized body for another, a phenomenon known before 600 B.C.

Atomic magnetic moments arise when the electrons of an atom possess a net magnetic moment as a result of their angular momentum. The combined effect of the atomic magnetic moments can give rise to a relatively large magnetization, or magnetic moment per unit volume, for a given applied field. Above the Curie temperature, a ferromagnetic substance behaves as if it were paramagnetic: Its susceptibility approaches the Curie-Weiss law. The Curie temperature marks a transition between order and disorder of the alignment of the atomic magnetic moments. Some materials having atoms with unequal moments exhibit a special form of ferromagnetism below the Curie temperature called ferrimagnetism. *See* CURIE TEMPERATURE; CURIE-WEISS LAW; ELECTRON SPIN; FERRIMAGNETISM; MAGNETIC SUSCEPTIBILITY; PARAMAGNETISM.

The characteristic property of a ferromagnet is that, below the Curie temperature, it can possess a spontaneous magnetization in the absence of an applied magnetic field. Upon application of a weak magnetic field, the magnetization increases rapidly to a high value called the saturation magnetization, which is in general a function of temperature. For typical ferromagnetic materials, their saturation magnetizations, and Curie temperatures, *See* MAGNETIZATION.

The tasks of a theory of ferromagnetism are to account for the spontaneous magnetization below the Curie point, the temperature dependence of the saturation magnetization, and the nature of the magnetization process, or magnetization curve.

Weiss theory. The Weiss molecular field theory of ferromagnetism (P. Weiss, 1907) represents the first realistic attempt to account for the properties of a ferromagnet. This theory rests on two hypotheses.

1. Below the Curie point, a ferromagnetic substance is composed of small, spontaneously magnetized regions called domains. The total magnetic moment of the material is the vector sum of the magnetic moments of the individual domains. It is now known that these assumed domains really exist and are usually between 0.01 and 0.1 cm wide.

2. Each domain is spontaneously magnetized because a strong internal magnetic interaction tends to align the individual atomic magnetic moments within the domain.

The consequence of these assumptions is that, while each domain is spontaneously magnetized, the directions of magnetization of the domains do not coincide; therefore the overall magnetization of the sample may be much smaller than if it were composed of a single domain. Application of a relatively weak field of the order of 10–100 kiloamperes/m or 100–1000 oersteds (and often very much less) is sufficient to align the directions of magnetization of the domains, thereby achieving a large magnetization.

The second hypothesis of the Weiss theory leads to the existence of a Curie temperature below which a domain may be spontaneously magnetized in the absence of an applied magnetic field. This comes about in the following way. If the domain is spontaneously magnetized, there must be some sort of interaction between the atomic magnetic moments which tends to align them. Otherwise the domain would behave paramagnetically. The average strength of this interaction may be represented by an internal magnetic field, the Weiss molecular field, which is proportional to the magnetization of the domain. Thus the effective field acting on any atomic magnetic moment within the domain may be written as Eq. (1), where H_0 is an externally applied magnetic field and

$$H = H_0 + \lambda M \tag{1}$$

λM (M = magnetization) is the Weiss molecular field. For iron $M \simeq 135$ kA/m (1700 Oe) and $\lambda \simeq 5000$, so that $\lambda M \simeq 10^9$ A/m (10^7 Oe), a huge value. It is relatively easy to deduce the magnetic susceptibility above the Curie point since it is known that the ferromagnetic substance behaves like a paramagnet above the Curie temperature. If the Curie law holds, Eq. (2) is valid, where C is the Curie constant.

$$\frac{M}{H} = \frac{C}{T} \tag{2}$$

In the Curie law, H is taken to be the effective magnetic field acting on an atomic magnetic moment, Eq. (1). From Eqs. (1) and (2), Eq. (3) is obtained.

$$\frac{M}{H + \lambda M} = \frac{C}{T} \tag{3}$$

The susceptibility χ is the magnetization per unit applied field H_0, so that from Eq. (3), in the electromagnetic system of units, Eq. (4) is derived.

$$\chi = \frac{M}{H_0} = \frac{C}{T - C\lambda} = \frac{C}{T - T_c} \tag{4}$$

The Curie temperature is defined by $T_c = C\lambda$, and the susceptibility follows the Curie-Weiss law above the Curie temperature. The form of the Curie-Weiss law leads to a nonzero magnetization when $H_0 = 0$ at T_c.

Below the Curie point, where the Curie-Weiss law breaks down, the Weiss theory leads to a spontaneous magnetization and also predicts the temperature dependence of the saturation magnetization. This comes about if it is taken into account that the Curie law is only an approximation valid for weak fields or high temperatures; it does not allow for saturation effects. The correct quantum-mechanical expression to replace Eq. (2) is given by Eqs. (5) and (6), where N is the number of atoms per unit volume,

$$M = NJg\mu_B B_J(a^*) \tag{5}$$

$$a^* = \frac{g\mu_B(H_0 + \lambda M)J}{kT} \tag{6}$$

each with angular momentum quantum number J, g is the spectroscopic splitting factor (the measure of the energy level splittings), μ_B is the Bohr magneton, k is Boltzmann's constant, T is the absolute temperature, and $B_J(a^*)$ is the Brillouin function. For $T < T_c$, the value of the magnetization for an applied field of H_0 is obtained by solving Eqs. (5) and (6) simultaneously. In particular, the spontaneous magnetization is obtained by setting $H_0 = 0$ in Eq. (6). *See* GYROMAGNETIC EFFECT; GYROMAGNETIC RATIO; MAGNETON.

The solution for the spontaneous magnetization ($H_0 = 0$) may be obtained graphically for any temperature T, as shown in **Fig. 1**. The broken line of Fig. 1 represents the line $M = a^*kT_c/g\mu_B J\lambda$ and is the line for the largest value of T for which there is a solution. This is the Curie point, and for $T > T_c$ there is no spontaneous magnetization. The value of T_c obtained in this way [the slope of Eq. (5) at $a^* = 0$] is $T_c = Ng^2\mu_B{}^2J(J + 1)\lambda/3k$, a result which is consistent with the earlier definition $T_c = C\lambda$. The results obtained from Fig. 1 are plotted as a function of temperature in **Fig. 2**. The value of J chosen is $^1/_2$.

The plot of Fig. 2 is in overall general agreement with the experimental results for iron, nickel, and cobalt. The low-temperature behavior, however, is better described by magnon theory. *See* MAGNON.

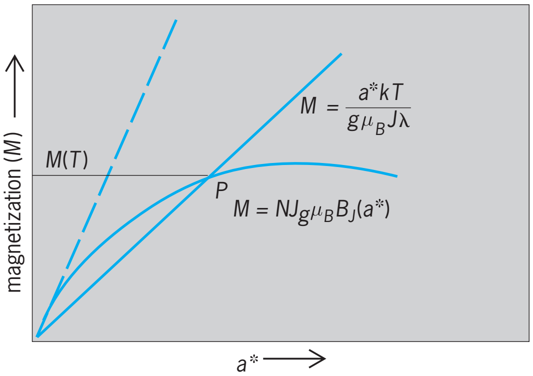


Fig. 1. Spontaneous magnetization below the Curie point. Intersection *P* determines *M*(*T*).

Heisenberg theory. The Heisenberg theory of ferromagnetism (W. Heisenberg, 1928) treats the origin of the Weiss molecular field on an atomic basis. It may be remarked at the outset that ordinary dipole-dipole interactions among atomic magnetic moments are much too small to account for the Weiss field. The foundation of the Heisenberg theory is the Pauli exclusion principle. The electrons providing the atomic magnetic moments are described by wave functions containing both space coordinates and spin angular momentum coordinates. The Pauli principle demands that the electron wave functions change sign (that is, be antisymmetric) when the space and spin coordinates of a pair of electrons are switched, or exchanged. As a consequence, the energy of the electrons depends on their relative spin orientation. Heisenberg realized that the essential

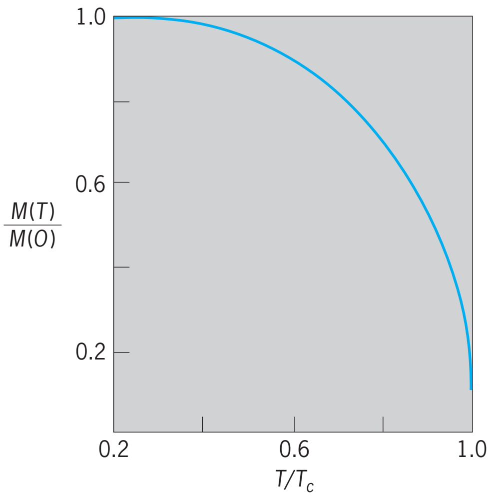


Fig. 2. Temperature dependence of spontaneous magnetization for *J* = $^1/_2$.

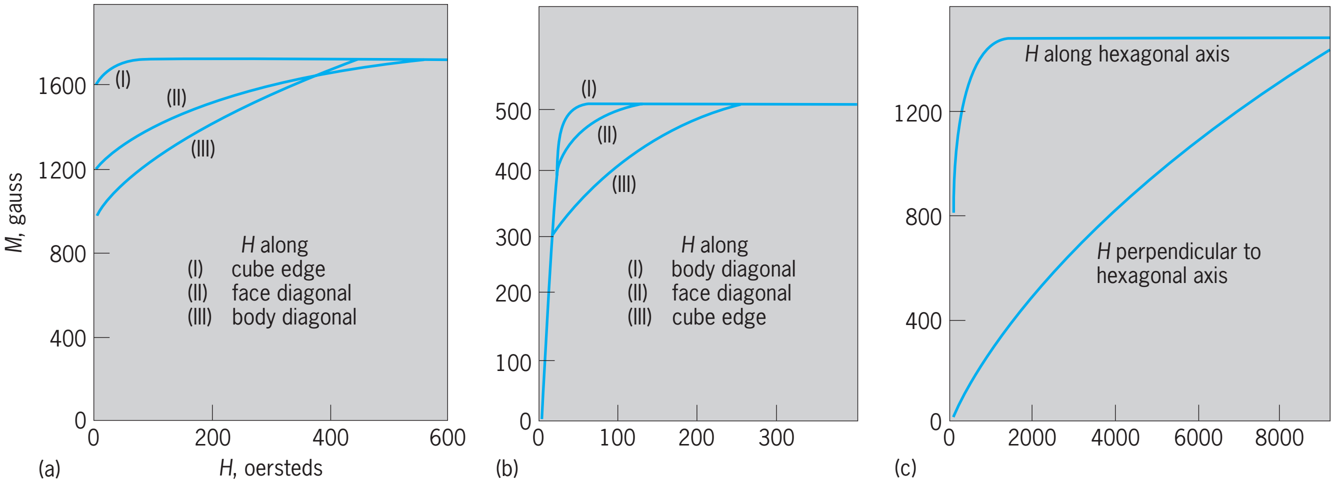


Fig. 3. Magnetization curves, for various directions of applied field *H*, indicated for single crystals (*a*) iron, (*b*) nickel, and (*c*) cobalt. 1 gauss = 10^{-4} tesla; 1 oersted = 79.6 A/m.

physics involves treating the electrostatic Coulomb repulsion between electrons within the constraint of the Pauli exclusion principle. He showed that the situation can be viewed in terms of an effective interaction between the electron spins, the exchange interaction energy, which had first been derived by P. A. M. Dirac. In Eq. (7), $\mathbf{S}_i$ and $\mathbf{S}_j$ are the spin angular momentum vectors of the two electrons i and j, and J_{ij} is the so-called exchange integral between the electrons. The exchange integral decreases rapidly with distance between the electrons and depends in a complicated way upon the spatial distribution (wave function) of the electrons. It is extremely difficult to compute. If the exchange integral is positive, the parallel arrangement is favored, and if J is large enough, ferromagnetism should result. If J is large and negative, antiferromagnetism or ferrimagnetism supposedly arises. The order of magnitude of J given by $J \simeq kT_c \simeq 10^{-13}$ erg. *See* EXCLUSION PRINCIPLE.

$$\text{Exchange energy} = -2J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j \qquad (7)$$

There seems to be no question that the Heisenberg theory correctly accounts for the tendency of electrons in the same ferromagnetic atom to exhibit parallel spins. However, whether or not it leads to the correct explanation of the interatomic alignment of spins is still a subject of controversy. In insulators it usually proves possible to express the coupling between atomic spins by Eq. (7), J_{ij} is interpreted as an effective exchange integral. In metals the problem is much more complicated. There is little doubt, however, that the basic Heisenberg idea is correct and the molecular field arises from the interplay between electrical forces and the effects of Pauli exclusion.

Magnetocrystalline anisotropy energy. This accounts for the experimental fact that ferromagnets tend to magnetize along certain crystallographic axes, called directions of easy magnetization. For example, a single crystal of iron, which is made up of a cubic array of iron atoms, tends to magnetize in the directions of the cube edges. At room temperature the magnetization of iron can be moved into a hard direction along a body diagonal only if an external magnetic field greater than about 30 kA/m (400 Oe) is applied in that direction. This can be seen in **Fig. 3**, which shows single-crystal magnetization curves for various directions of applied field *H* in iron, nickel, and cobalt.

N. S. Akulov showed (1929–1931) that the anisotropy energy U_A could be expressed conveniently in an ascending power series of the direction cosines between the magnetization and the crystal axes. For cubic crystals the lowest-order terms take the form of Eq. (8), where α_1, α_2, and α_3 are direction cosines with respect to the three cube edges, and K_1 and K_2 are temperature-dependent parameters characteristic of the material, called anisotropy constants. In general, $|K_1| > |K_2|$, and further terms are unnecessary. In iron, K_1 is positive, and therefore U_A is a minimum when any single direction cosine $\alpha_i = 1$ and the other $\alpha_j = 0$. That is, the cube edges are easy directions. In nickel, K_1 is negative and hence the body diagonals are easy directions (Fig. 3).

$$U_A = K_1({\alpha_1}^2{\alpha_2}^2 + {\alpha_2}^2{\alpha_3}^2 + {\alpha_3}^2{\alpha_1}^2) + K_2({\alpha_1}^2{\alpha_2}^2{\alpha_3}^2) \qquad (8)$$

For crystals of other than cubic symmetry the energy U_A must be expressed in a form which is different from Eq. (8) and appropriate to the particular symmetry.

Anisotropy constants can be determined from (1) analysis of magnetization curves, (2) the torque on single crystals in a large applied field, and (3) single-crystal ferromagnetic resonance. *See* MAGNETIC RESONANCE.

The Heisenberg exchange energy, Eq. (7), is isotropic and cannot account for the observed anisotropy, which probably has its origin in a complicated interplay of spin-orbit coupling, crystalline electric fields, and overlap of orbital wave functions. The anisotropy energy also depends on the state of strain of the crystal, giving rise to magnetostriction,

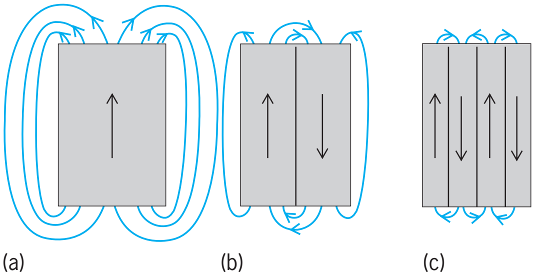


Fig. 4. Lowering of magnetic field energy by domains. (*a*) Lines of force for a single domain. (*b*) Shortening of lines of force by division into two domains. (*c*) Reduction of field energy by further subdivision.

that is, changes in length of a substance when it is magnetized. *See* MAGNETOSTRICTION.

Ferromagnetic domains. Small regions of spontaneous magnetization, formed at temperatures below the Curie point, are known as domains. As shown in **Fig. 4**, domains originate in order to lower the magnetic energy. In Fig. 4*b* it is shown that two domains will reduce the extent of the external magnetic field, since the magnetic lines of force are shortened. On further subdivision, as in Fig. 4*c*, this field is still further reduced.

Another way to describe the energy reduction is to note that the interior demagnetizing fields, coming from surface poles, are much smaller in the long, thin domains of Fig. 4*c* than in the "fat" domain of 4*a*.

The question arises as to how long this subdivision process continues. With each subdivision there is a decrease in field energy, but there is also an increase in Heisenberg exchange energy, since more and more magnetic moments are aligning antiparallel. Finally a state is reached in which further subdivision would cause a greater increase in exchange energy than decrease in field energy, and the ferromagnet will assume this state of minimum total energy.

Bloch wall. Also because of exchange energy, the reversal of magnetization between domains does not occur abruptly but takes place gradually through a transition zone called the Bloch wall (**Fig. 5**). To understand the reason for this wall, consider the exchange energy involved. Let the angle between the magnetization of neighbor planes be ϕ and assume that, as is usual, the exchange energy is appreciable only between neighbor atoms. According to Eq. (7), the exchange energy between atoms on neighbor planes varies as $-\cos\phi$, or as a constant $+\phi^2/2$ if ϕ is small. If total reversal occurs in N planes, $\phi = \pi/N$, and the total exchange energy of the wall is given by expression (9). This equation shows that the total ex-

$$U_E \text{ (wall)} \propto \text{constant} + N\left(\frac{\pi}{N}\right)^2 \qquad (9)$$

change energy decreases as the number of transition planes increases; that is, the exchange energy tends to make the domain wall thicker.

The question now arises as to why N should not continue to increase until the entire crystal becomes a single Bloch wall. This does not happen because the decrease in exchange energy is accompanied by an increase in anisotropy energy. Many of the intermediary planes of Fig. 5*b* must of necessity have their magnetization along hard directions. The larger the value of N, the greater must be the number of such planes. The thickness of the wall can be determined by finding that value of N which makes the sum of exchange and anisotropy energies a minimum. In iron the wall is ~50 nanometers thick and has the total energy ~2 ergs/cm^2.

Domain arrangement. The orientation of domains in a crystal is determined primarily by the need to minimize the magnetic energy (Fig. 4). It is possible to eliminate all surface magnetic poles by forming flux-closure domains (**Fig. 6**). Here the normal component of magnetization is continuous across all domain boundaries. The demagnetizing fields are zero everywhere, except for the trivial effect of surface poles in the Bloch walls. In a uniaxial crystal, that is, a crystal with a single easy direction, an arrangement as shown in Fig. 6*b* will be preferred since it has a lower density of magnetization normal to the easy direction, or in the hard direction. Even in cubic crystals, in which all directions of magnetization in Fig. 6 may be easy, Fig. 6*b* will be preferred because of magnetostriction. In iron, for example, each domain increases in length along the direction of magnetization by a fraction ~2 × 10^{-5}. Thus the domains of Fig. 6 can be fitted smoothly together only by straining them elastically against this magnetostriction, and the required elastic strain energy will be smaller in Fig. 6*b* than in Fig. 6*a*.

In polycrystals the domain structure is much more complicated, depending upon such variables as grain

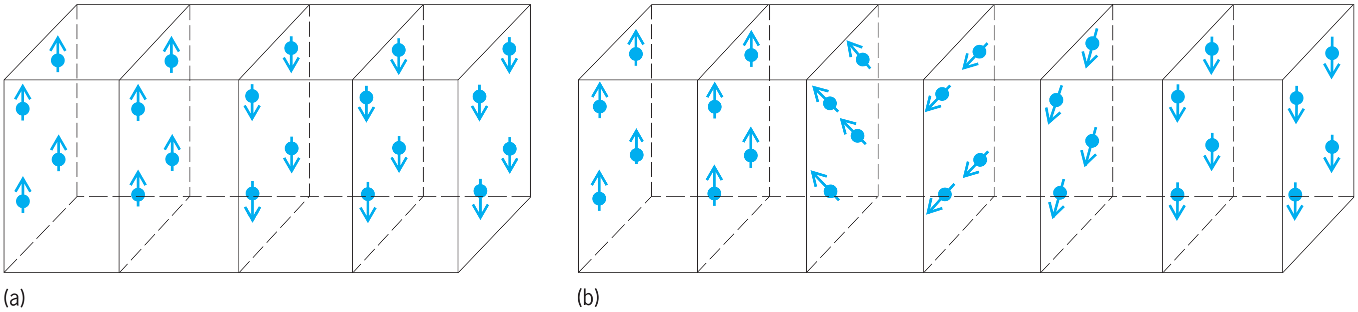


Fig. 5. Lowering of exchange energy by the transition zone known as Bloch wall. The reversal of magnetization between domains (*a*) does not take place abruptly but (*b*) by degrees.

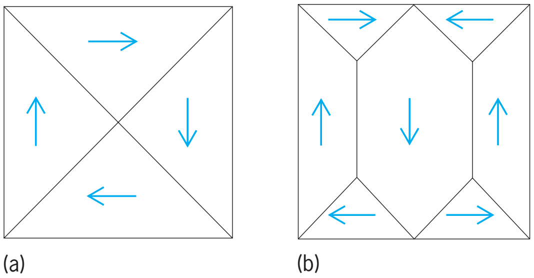


Fig. 6. Flux-closure domains. (*a*) Large domains at right angles. (*b*) Reduction in their size, causing reduction of anisotropy energy of uniaxial crystals or of strain energy of cubic crystals.

orientation and grain boundaries. It is possible, however, for domains to cross grain boundaries.

On minimizing the total contributions from (1) magnetic, or demagnetizing, energy, (2) anisotropy, (3) magnetostriction, (4) elastic strain, and (5) Bloch wall energy, it is found that, depending upon the composition and shape of the crystal, the theoretical domain thickness should vary from about 0.001 to 0.1 cm.

Bitter powder patterns. Direct experimental evidence of the existence of domains is furnished by Bitter powder patterns. First a surface is prepared by careful polishing with fine abrasive, followed (for metallic samples) by electropolish. Then a drop of a colloidal suspension of magnetic Fe_2O_3 particles is placed on the surface and covered by a microscope cover glass. These particles gather at surface regions where the magnetic field is largest and can be observed easily with a microscope. It is seen from Fig. 5 that the surface field is largest at the center of the Bloch wall; that is, lines of surface poles will demarcate the wall. Thus the colloidal particles concentrate along domain boundaries, as shown in **Fig. 7**. If the surface deviates by even a few degrees of angle from a simple lattice plane, the pattern becomes very complicated. The beginnings of this may be seen in the "fir tree" patterns at the lower right of Fig. 7.

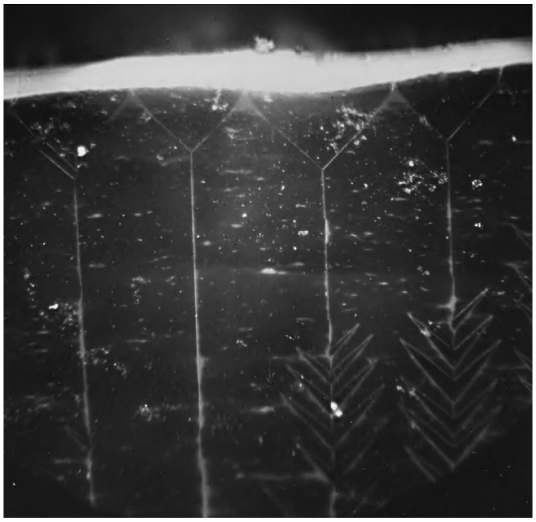


Fig. 7. Bitter powder patterns on a (100) surface of silicon-iron. (*Photograph by H. J. Williams*)

The powder patterns only delineate the boundaries and do not themselves show the direction of magnetization. This direction may, however, be inferred from a study of how particles gather around a scratch made on the surface, coupled with observation of domain growth when the sample is placed in a magnetic field.

The Kerr magnetooptic effect (rotation of the plane of polarization of light reflected from a magnetic surface) has also been used to study domains. J. F. Dillon demonstrated that single-crystal slabs of ferromagnetic yttrium iron garnet (YIG) are transparent to visible light and that the Faraday effect (rotation of the plane of polarization of light transmitted along a magnetic field) can be used to observe domains. The effect is very striking, and the growth and diminution of domains are readily visible. *See* FARADAY EFFECT; KERR EFFECT.

It is also possible to observe domains by means of transmission or reflection of electrons in an electron microscope. This technique, and also the optical techniques mentioned above, are particularly useful if the Bloch walls are extremely thick (because of small magnetocrystalline anisotropy) and the resultant Bitter powder patterns very blurred.

Bloch wall motion. This accounts for the initial portion of the magnetization curve. As the wall passes through regions of crystal strain, or over a number of foreign atoms, it may suddenly go over a potential energy hump and into a minimum. There will be a small, almost discontinuous jump in the magnetization, easily detected by a search coil (a device for measuring change of flux density) and amplifier. This is called the Barkhausen effect. It was once thought erroneously that this effect comes from a sudden, complete reversal of magnetization within a domain. However, H. J. Williams and W. Shockley showed that many Barkhausen jumps accompany the motion of a single domain wall. *See* BARKHAUSEN EFFECT; DOMAIN (ELECTRICITY AND MAGNETISM).

Soft magnetic materials. Materials easily magnetized and demagnetized are called soft; these are used in alternating-current machinery. The problem of making cheap soft materials is complicated by the fact that readily fabricated metals usually have many crystalline boundaries and crystal grains oriented in many directions. In such metals the magnetization process is accompanied by much irreversible Bloch wall motion and by much rotation against anisotropy, which is usually irreversible.

The ideal cheap soft material would be an iron alloy fabricated by some inexpensive technique which results in all crystal grains being oriented in the same or nearly the same direction. Various complicated rolling and annealing methods have been discovered in the continued search for better grain-oriented or "cube-textured" steels.

In some situations the cost of fabrication is secondary, and alloys with the smallest possible crystalline anisotropy and magnetostriction are demanded. The first such need appears to have been for the inductive loading of submarine telegraph cables. This gave rise to the perfection of permalloy, an alloy of 21.5% iron and 78.5% nickel. The constant K_1 of Eq. (8) takes opposite signs in pure iron and pure nickel, and at the permalloy composition the resultant anisotropy goes to zero. More precisely, the composition for zero anisotropy is not quite the same as for zero magnetostriction, and permalloy represents a compromise. Addition of a third element, generally molybdenum, can simultaneously drive anisotropy and magnetostriction to zero and, as a bonus, decrease resistivity.

Hard magnetic materials. Materials which neither magnetize nor demagnetize easily are called hard; these are used in permanent magnets. The properties of a hard magnet material are described by its B versus H hysteresis curve, where B is the magnetic induction and H the applied magnetic field. Important material parameters specified by this curve include the remanence B_r (the induction for zero applied field), the coercivity or coercive force H_c (the reverse applied field required to drive B to zero), and the maximum energy product $(BH)_{\max}$, the maximum product of B and H in the second quadrant of the hysteresis curve. The parameter $(BH)_{\max}$ is the most widely used single figure of merit for a permanent magnet material; it provides a measure of the magnetic field that a unit volume of the material can produce.

Lodestone was the first magnetically hard material known. Principally composed of the ore magnetite (Fe_3O_4) it was long applied in navigational compasses. In the twentieth century, five classes of permanent-magnet materials have enjoyed technological importance. The magnet steels contain carbon, chromium, tungsten, or cobalt additives, serving to impede domain wall motion and thus to generate coercivity. These materials feature high remanence ($B_r \simeq 1$ tesla or 10 kilogauss) but low H_c and low $(BH)_{\max}$ (less than 8 kT/m^3 or 1 megagauss-oersted). Alnicos are aluminum-nickel-iron alloys containing finely dispersed, oriented, elongated particles precipitated by thermal treatment in a field. The coercivity arises from the shape of the particles, and $(BH)_{\max}$ can be as large as 80 kT/m^3 (10 MG · Oe). Hard ferrite magnets are based on the oxides $BaFe_{12}O_{19}$ and $SrFe_{12}O_{19}$. Their coercivity derives from the magnetocrystalline anisotropy, rather than the shape, of small, single-domain particles, and can exceed that of alnicos. Although $(BH)_{\max}$ is no larger than 40 kT/m^3 (5 MG · Oe), hard ferrite magnets are relatively inexpensive and are used in a great variety of commercial applications.

Substantial increases in $(BH)_{\max}$ have been afforded by rare earth–transition metal materials whose rare earth component provides huge magnetocrystalline anisotropy which can be translated into large coercivity in a practical magnet, while the magnetization arises chiefly from the transition metal component. Samarium-cobalt magnets based on the $SmCo_5$ or Sm_2Co_{17} intermetallic compounds can have coercivities of ~1.6 MA/m (20 kOe) and maximum energy products in the range of 160–240 kT/m^3 (20–30 MG · Oe). Magnets containing the more abundant elements neodymium and iron rather than samarium and cobalt have been developed from neodymium-iron-boron materials based on the ternary compound $Nd_2Fe_{14}B$. Such magnets have been prepared both by the powder metallurgy method used for ferrite and samarium-cobalt materials and by the rapid solidification technique of melt spinning in which molten alloy is directed onto a rapidly rotating substrate wheel. Neodymium-iron-boron magnets with $(BH)_{\max}$ approaching the enormous value of 400 kT/m^3 (50 MG · Oe) have been produced. Elihu Abrahams; Frederic Keffer; J. F. Herbst

Bibliography. B. Barbara, D. Grignoux, and C. Vettier, *Lectures on Modern Magnetism*, 1988; R. M. Bozorth, *Ferromagnetism*, 1994; S. Chikazumi and S. H. Charap, *Physics of Magnetism*, 1964, reprint 1978; B. D. Cullity, *Introduction to Magnetic Materials*, 1972; R. A. McCurrie (ed.), *The Structure and Properties of Ferromagnetic Materials*, 1994; D. C. Mattis, *Theory of Magnetism*, vol. 1, 2d ed., 1988, vol. 2, 1985; E. P. Wohlfarth (ed.), *Ferromagnetic Materials*, 3 vols., 1980–1982.

Ferry

A ship specifically configured for carrying passengers between two points. It permits persons to make their way from one place to another across a body of water, and it may carry vehicles, including commercial vehicles. A ferry is distinct from a cruise ship or a cargo ship. For a cruise ship the voyage itself is the destination, whereas with a ferry the journey's end point is the destination. While a ferry may carry cargo, in contrast to a cargo ship, this cargo is contained in a commercial vehicle accompanied by a driver.

Technology. Ferries may range in appearance from ships that are outwardly indistinguishable from cruise ships to small craft indistinguishable from pleasure boats. Some ferries with overnight runs on the Baltic Sea carry well over a thousand passengers, housing them in luxurious cabins and offering casinos and theaters for entertainment. At the other end of the spectrum, many of the world's harbors are served by water taxis—small ferries which may carry as few as six persons for a 5–15-min run.

Ferries may range in speed from slow vessels of 10 knots (18.5 km/h) or less to ultrafast ferries with speeds exceeding 60 knots (110 km/h). They may carry passengers only, or they may carry a mixture of passengers and private automobiles, or passengers, cars, and trucks. Ferries may be characterized by their hull form and their propulsion plant.

Hull forms. Ship hull forms are characterized by the way they support the weight of the ship. This is depicted in the sustention triangle (**Fig. 1**). The vertices of this triangle depict the three modes employed to support a ship. The buoyant, or archimedean, support mode is the fundamental means of support of all ships and boats.

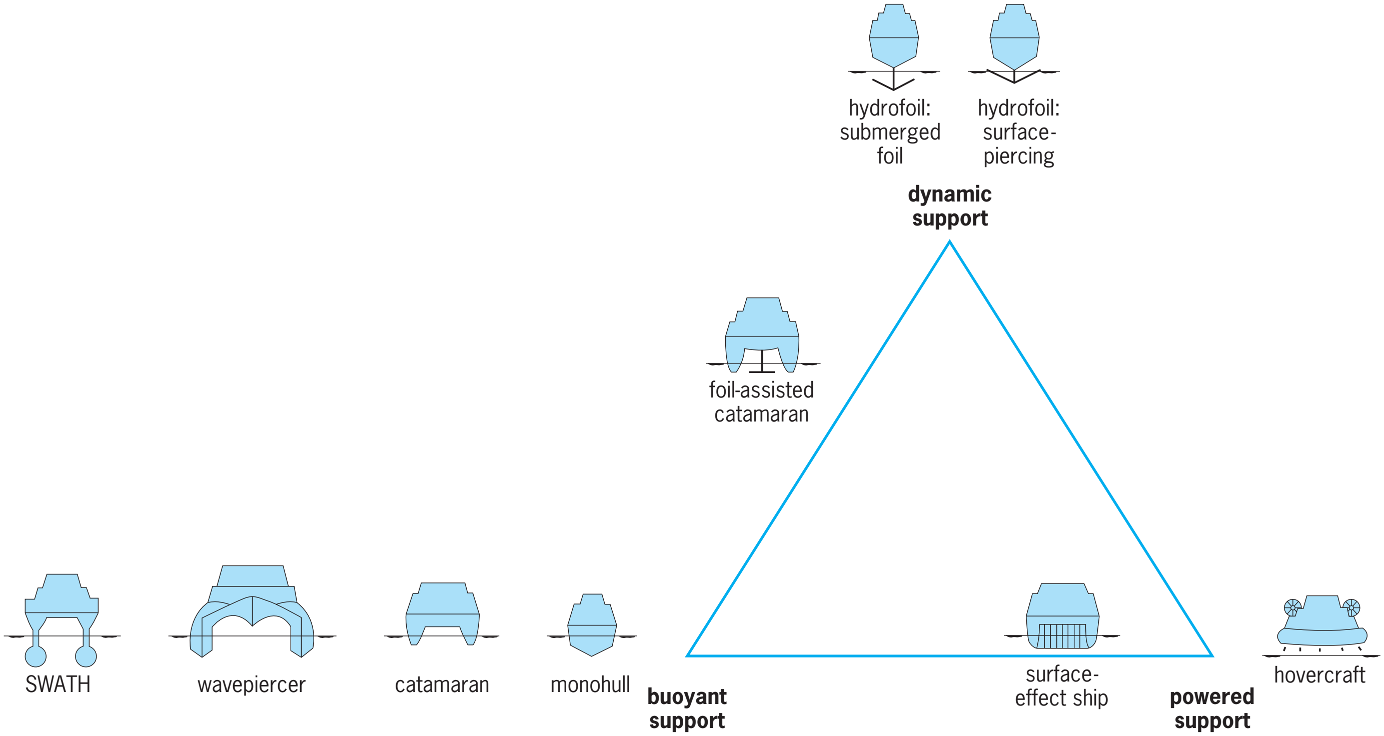


Fig. 1. **Sustention triangle depicting the three principal modes of supporting the weight of a ferry and the types of ships that employ them.**

Buoyant support. In buoyant support mode, a vessel's weight is borne by the upward force produced by the static displacement of the water in which it floats. The weight of the volume of water displaced is exactly equal to the weight of the ship. Any increase in the weight of the ship (for example, due to the loading of passengers or cargo) requires the ship to displace an increased volume of water, which it does by sinking to a slightly deeper draft. *See* ARCHIMEDES' PRINCIPLE; BUOYANCY; HYDROSTATICS.

The most common hull design for ferries is the displacement monohull ship form, which has one hull. Monohull ferries have been built in all sizes and speeds. Among monohulls a unique form is the double-ended ferry (**Fig. 2**). On short runs, it makes operational sense to build a ship which is identical fore and aft. For example, in use for river crossing, this ship simply shuttles from one dock to the other, putting one "bow" into the slip at one port, the other "bow" at the other port.

The catamaran hull is a very popular form, particularly for high-speed ferries (**Fig. 3**). A catamaran is a twin-hulled ship, having two hulls side by side. Because the two hulls are slender, a catamaran may have less wavemaking drag than a monohull of the same displacement. The two hulls are generally separated by a distance about equal to the beam of one hull. This results in a broad main deck area, which is more nearly square than the slender deck found on a monohull ship. Because of the broad expanse of the deck area, and because its low-aspect-ratio shape is easier to load with vehicles, the catamaran hull form has been very successfully incorporated into ferry service.

The drag (resistance to forward motion) of a displacement ship has three main components: air drag associated with moving the above-water portion of the ship, and two hydrodynamic drag components, skin friction and wavemaking drag. Skin friction is the drag of water sliding along the wetted surface of the hull.

Fig. 2. **Double-ended ferry. (*Washington State Ferries*)**

Fig. 3. **High-speed catamaran ferry. (*Hitachi Zosen*)**

When a ship moves through the water, it produces waves known as the ship's wake. Forming waves in the water surface requires a tremendous amount of energy. In order to move forward, a displacement ship must force aside a mass of water equal to the mass of the ship. Fast forward movement requires this substantial mass to be displaced rapidly, and the energy required grows. To a first order of magnitude, the power required to move a displacement-type hull will increase as the cube of the speed. *See* SHIP POWERING, MANEUVERING, AND SEAKEEPING.

The steep slope of the cubic speed-power relationship has driven engineers to seek ship hull forms with less drag. The quest for reduced drag has focused on finding means of support other than displacement, so that the ship no longer relies on thrusting aside a large mass of water.

Dynamic support. Shown in the sustention triangle (Fig. 1), dynamic support is produced by the movement of the ship and disappears when the vessel is at rest. The most common example of dynamic support is planing, the skimming of a craft across the surface of the water. In this case, the dynamic support is produced by the shape of the hull itself. Most small fast motor boats are planing craft.

A hydrofoil is a form of dynamic support craft which uses submerged winglike appendages to produce dynamic lift. These foils lift the main hull of the craft clear of the sea surface. The wavemaking portion of craft drag is thereby reduced; lifting the craft above the surface reduces the energy dissipated in wavemaking. The frictional component of drag, however, is undiminished, or may even be increased due to the increased surface area of the hydrofoils. Since the advantages of this type of craft come from the lifting of the ship above the water, it becomes important to reduce the basic weight of the craft. Lightweight materials and components are frequently employed on dynamically supported hulls. *See* HYDROFOIL CRAFT.

Powered support. The third vertex of the sustention triangle denotes powered support. In the marine world the only example is the air-cushion vehicle (ACV), including the related surface-effect ship (SES). With the air-cushion vehicle, an engine-driven fan produces aerostatic pressure beneath the bottom of the craft, lifting the rigid hull above the water surface. In this case there may be no contact between the ship's structure and the water, and thus there may be no frictional drag. The air pressure does, however, displace the water surface, and thus there is still wavemaking drag. Again, since the powered lift craft expends energy to raise the ship, it is in the designer's interest to reduce the weight of the ship.

A pure air-cushion vehicle has an air cushion underneath all parts of the ship—there is no contact between hard structure and the water. This produces some engineering challenges, the most obvious of which is that it is difficult to use underwater propulsors such as marine screw propellers. To overcome this, a hybrid type of vehicle was invented which is called a surface-effect ship in the United States. In the United Kingdom this concept is called a sidewall hovercraft. The surface-effect ship is a marriage of the catamaran and the air-cushion vehicle. A pair of slender catamaran-like hulls have an air cushion between them. Surface-effect ships are well suited to very high speeds (as much as 80 knots or 150 km/h). A few surface-effect ship ferries exist. *See* AIR-CUSHION VEHICLE.

Propulsion systems. Ferry propulsion systems depend upon the speed and hull form of the ship. The most common propulsion system is the conventional screw propeller. This is a submerged helical screw at the stern of the ship, driven by the ship's engines. Double-ended ferries require some form of reversible or double-ended propulsion system. Large double-ended ferries, such as those operated by the State of Washington and City of New York, have rudders and propellers at each end, sometimes having separate engines for each end. Other, much smaller double-ended ferries use a form of reversible propulsion, such as an azimuthing drive located amidships. This drive is steered 180° when the vessel reverses course. *See* PROPELLER (MARINE CRAFT).

High speed ferries—none of which are double-ended—face special propulsion issues. Whereas the diesel engine is the prime mover of choice for low-speed ferries, fast ferries demand much higher power. The higher weight sensitivity of a fast ferry frequently leads to use of gas turbine engines because of their light weight. *See* MARINE ENGINE.

Virtually all low-speed ferries and many high-speed ferries are propeller-driven. However, at speeds above about 40 knots (75 km/h) the hydrodynamic efficiency of the waterjet propulsor becomes superior to that of a propeller. As a consequence, almost all the fastest high-speed ferries use waterjet propulsion (**Fig. 4**).

Operation and economics. The condition which makes ferry service attractive is simple: a population divided by water. This might be, for example, a large city which is penetrated by a harbor or lake. In such a case, persons wishing to go to a point directly opposite them across the lake or harbor have the choice of taking a long land route or a considerably shorter water route. This same principle may

Fig. 4. Large waterjet propulsors used in high-speed catamaran ferry. (*Hitachi Zosen*)

apply on a much large scale to, for example, the Baltic or Mediterranean seas. It is, in fact, possible to drive from Italy to Africa, but the route is ridiculously long, and the maritime route is greatly preferred. Island communities are other obvious users of ferries.

It is important that ferries be able to load and unload rapidly. This is best accomplished by maximizing the number of lanes of traffic that can move in parallel, and eliminating obstacles in those lanes. Thus a fast-loading ferry may have all vehicle lanes continuous across the shore ramps, and may be capable of loading all vehicles straight-ahead, without need for any turning, backing, or other maneuvering. Similar requirements apply to pedestrian loading, wherein it is important to have wide lanes and eliminate bottlenecks created by narrow doorways, gate agents, and so forth. The rapidity of load and unload is a major factor in the economic success of a ferry, since a ferry which loads inefficiently may not offer any time savings compared to nonferry alternatives.

Laws and regulations. Most nations have federal maritime safety agencies regulating the operation of ferries within their national waters. In the United States, this role is performed by the Coast Guard. Ferries which run on international routes generally also fall under the authority of the International Maritime Organisation.

A significant secondary role is played by a group of private commercial entities known as classification societies. These societies were originally developed by insurers to assess the safety and insurability of a ship. In the centuries of their existence, they have published a large body of guidance on design, construction, operation, and maintenance of ferries. These guidelines have been adopted by some flag states and given the stature of de facto national law.

The largest international classification societies are the American Bureau of Shipping, Lloyds Register, and Det Norske Veritas. There are many others, and the selection of a classification society is a commercial decision made by the ship owner. *See* SHIP DESIGN.

Chris B. McKesson

Bibliography. E. V. Lewis (ed.), *Principles of Naval Architecture*, 2d ed., 3 vols., Society of Naval Architects and Marine Engineers, Jersey City, 1988; *Rules for Building and Classing Steel Vessels*, American Bureau of Shipping, Houston, 1998; R. Taggurt (ed.), *Ship Design and Construction*, 3d ed., Society of Naval Architects and Marine Engineers, Jersey City, 1980; Title 46, *US Code of Federal Regulations* (46 CFR).

Fertilization (animal)

The fusion of a female gamete (egg) with a male gamete (sperm) to form a single cell, the zygote. Gametes are haploid because their nuclei contain only one set of chromosomes. In contrast, most animal cells are diploid since their nuclei contain two sets of chromosomes, one set derived from the egg and the other from the sperm. The reduction from the diploid to the haploid number occurs in a specialized set of two cell divisions, known as meiosis, during which gametes are formed in the gonads of an adult. Fertilization restores the diploid chromosome number. In the human, normal eggs and sperm contain 23 chromosomes each, whereas the zygote and most of its daughter cells contain 23 pairs, or 46, chromosomes per cell nucleus. *See* GAMETOGENESIS; MEIOSIS; ZYGOTE.

Structure of eggs and sperm. Eggs and sperm contain the standard components of cells, including a nucleus and an outer membrane, the plasma membrane, which is a double layer of phospholipids with embedded proteins. However, each gamete also has specialized organelles for its particular function. The egg is the larger gamete. It sustains the embryo until the embryo can obtain nutrients from outside. This stage is reached quickly by the mammalian embryo when it implants in the uterus, so that the mammalian egg can be very small, typically about 0.1 mm in diameter. Other animal eggs are much bigger, as they contain storage materials that support embryonic development to more advanced stages. The mammalian egg is surrounded by the zona pellucida (or zona), a transparent envelope consisting of glycoproteins (**Fig. 1**). Other animal eggs, which are often released into the external environment, have additional protective layers. *See* CELL (BIOLOGY); OVUM.

The sperm's function is to find an egg of the same species and fuse with it. A mammalian sperm is sleek, with a head containing the highly condensed nucleus, a midpiece with mitochondria serving as the sperm's power plants, and a whiplike tail. Capping the nucleus is the acrosome, a bag of enzymes that will be released when the sperm penetrates the zona (**Fig. 2**). *See* SPERM CELL.

Attraction of sperm to eggs. In many animal species, sperm are attracted by chemicals released from the egg or—in mammals—from the ovarian follicle at ovulation. Evidently, the sperm can sense the attractant and swim to where the attractant is most concentrated. The sperm propels itself by the beat of its tail. In mammals, sperm need to move from

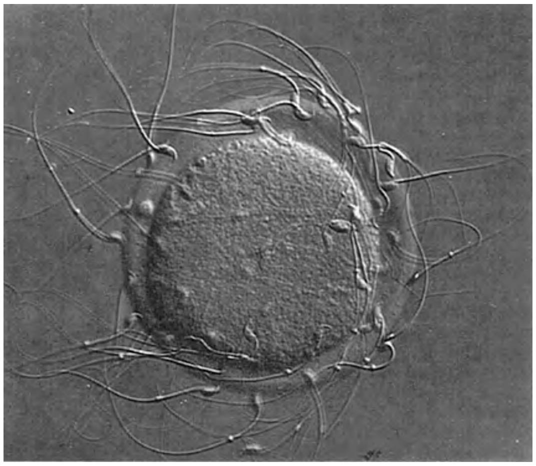


Fig. 1. Photomicrograph of mouse sperm bound to zona pellucida of an unfertilized egg under laboratory conditions (in vitro).

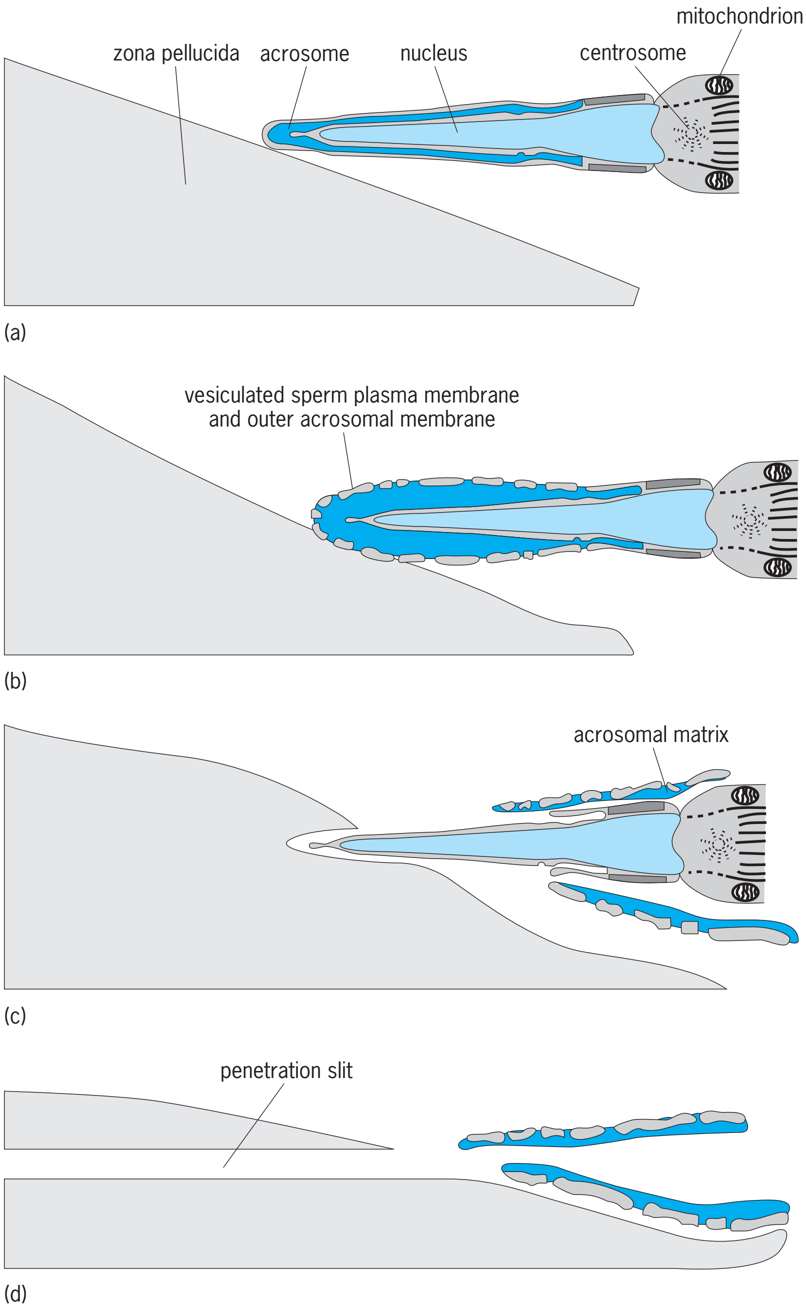


Fig. 2. Mammalian sperm (*a*) adhering to zona pellucida, (*b*, *c*) undergoing the acrosome reaction, and (*c*, *d*) penetrating the zona pellucida on its way to the egg cell proper. (*Modified from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

the vagina to the site of fertilization in the upper third of the oviduct. Although this travel is facilitated by movements of the female genital tract, only a few per million ejaculated sperm reach their destination. In humans, both egg and sperm remain viable in the oviduct and capable of fertilization for about 3 days. *See* REPRODUCTION (ANIMAL); REPRODUCTIVE SYSTEM.

Sperm–egg adhesion. An important step in the fertilization process is the adhesion of the sperm to the zona (or equivalent egg envelope in other animals) [Fig. 1 and Fig. 2*a*]. This step guarantees much of the species specificity of fertilization because sperm adhere selectively to eggs of the same species. This selectivity is mediated by matching molecules on the head of the sperm and the zona. Sperm–egg adhesion is associated with an event in the sperm known as the acrosome reaction (Fig. 2*b* and *c*). The sperm plasma membrane and the outer acrosomal membrane fuse at many points, breaking up the two membranes into many vesicles, so that the sperm tip is now covered by the inner acrosomal membrane. Lytic enzymes released from the acrosome, along with the sperm's propulsive action, help it to penetrate the zona (Fig. 2*c* and *d*).

Gamete fusion. In the defining event of fertilization, the plasma membranes of egg and sperm fuse to generate one membrane covering the contents of both gametes. This membrane fusion depends on a sperm membrane protein, Izumo, named after a Japanese marriage shrine. The egg now actively engulfs the contents of the sperm and draws them inside. The zygote then contains two haploid nuclei, called pronuclei. The female pronucleus originates when the egg completes its meiotic divisions. The male pronucleus is formed by decondensing the sperm nucleus. The sperm also contributes the centrosome, a pair of cylindrical bodies that had been located at the base of its tail. The centrosome duplicates, and each daughter centrosome forms one pole of a mitotic spindle. The spindle incorporates the chromosomes from both pronuclei and organizes the division of the zygote in two embryonic cells, or blastomeres. Both blastomeres will be diploid, combining the genetic information encoded in the chromosomes of egg and sperm. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY).

Egg activation. Gamete fusion triggers a rapid series of events collectively called egg activation. Within seconds, the electric potential across the zygote's plasma membrane changes from negative on the cytoplasmic side (resting potential) to positive (fertilization potential). Most of this change occurs by admitting sodium ions from outside the egg to the inside. A minute later, calcium ions are discharged from stores in the egg cytoplasm in a wave spreading from the sperm entry point over the entire egg. The released calcium activates regulatory proteins, which in turn trigger an acceleration of cell metabolism and the resumption of the cell division cycle, which had been arrested at the end of meiosis. *See* BIOPOTENTIALS AND IONIC CURRENTS; CELL CYCLE; CELL DIVISION.

Blocks to polyspermy. It is critical to allow only one sperm to fertilize an egg. Entry of more than one sperm into the egg cytoplasm, an event called polyspermy, would be disastrous because multiple introduced centrosomes would set up competing mitotic spindles, which would generate blastomeres with incomplete sets of chromosomes. Two complementary mechanisms to prevent polyspermy are known as the fast block and the slow block. The fast block is triggered by the fertilization potential, which involves a change in sperm receptors on the egg cell membrane and the release of calcium from the egg. The fertilization reaction does not allow any additional sperm to fuse with the egg. The fast block is instantaneous, but it lasts only for about a minute, after which the fertilization potential disappears. By that time, a slow but permanent block has been triggered by the calcium wave that passes over the egg. It triggers a dramatic change known as the cortical reaction because it occurs at the egg surface (cortex). Thousands of membrane-bound vesicles (cortical granules) fuse with the zygote's plasma membrane, thus shedding their contents into the space under the zona. These contents include enzymes that modify the zona glycoproteins so that already adhering sperm are released and newly arriving sperm do not adhere.

Klaus Kalthoff

Bibliography. K. Kalthoff, *Analysis of Biological Development*, 2d ed., McGraw-Hill, Boston, 2001; P. M. Wassarman (ed.), *Elements of Mammalian Fertilization*, vols. 1 and 2, CRC Press, 1991.

Fertilizer

Materials added to the soil, or applied directly to crop foliage, to supply elements needed for plant nutrition. These materials may be in the form of solids, semisolids, slurry suspensions, pure liquids, aqueous solutions, or gases.

The chemical elements nitrogren, phosphorus, and potassium are the macronutrients, or primary fertilizer elements, which are required in greatest quantity. Sulfur, calcium, and magnesium, called secondary elements, are also necessary to the health and growth of vegetation, but they are required in lesser amounts compared to the macronutrients. The other elements of agronomic importance, provided for plant ingestion in small (or trace) amounts, include boron, cobalt, copper, iron, manganese, molybdenum, and zinc. These minor elements are called micronutrients. All these fertilizer elements, along with other chemical elements, occur naturally in agricultural soils in varying concentrations and mineral compositions which may or may not be in forms readily accessible to root systems of plants. The addition of fertilizer to soils used for the production of commercial crops is necessary to correct natural deficiencies and to replace the components absorbed by the crops in their growth.

Plants also show toxic responses to certain elements (including some of the micronutrients) naturally present in some soils. Much laboratory and field testing has been done to develop an understanding of crop sensitivity and response to chemical elements in their various natural combined forms. Further research has led to methods of soil adjustment to enhance positive response and to suppress toxicity. Advances in sensitive analytical chemical techniques have done much to extend this knowledge. Plant species vary in sensitivity, and in some cases boundaries of the desirable operating concentration region between deficiency and toxicity are difficult to define. It is also possible that alternation of other soil components may shift availability of an in-place micronutrient to create a more favorable, or less favorable, condition without changing its concentration in the soil.

Commercial Fertilizers

The major effort of the chemical process industry, in regard to fertilizer, is directed to the production of compounds of the primary fertilizer elements. Crop requirements of fertilizer components could be satisfied by the spreading of individual materials for each element deficient in the soil. However, economy favors the single application of a balanced mixture that satisfies all nutritional needs of a crop. Since each primary element can be supplied in any number of different chemical compounds, the practice has given rise to a great diversity of formulations. Liquids, slurry suspensions, solids, and solid blends are marketed in a wide array of compositions to suit the needs of specific crops and the specific soils.

Formulas. The compositions of fertilizer mixtures, in terms of the primary fertilizer elements, are identified by an N-P-K code: N denotes elemental nitrogen; P denotes the anhydride of phosphoric acid (P_2O_5); K denotes the oxide of potassium (K_2O). All are expressed numerically in percentage composition, or units of 20 lb each per short ton (10 kg per metric ton) of finished fertilizer as packaged. Formula 8-32-16 thus contains a mixture aggregating 8 wt % N in some form of nitrogen compounds, 32 wt % P_2O_5 in some form of phosphates, and 16 wt % K_2O in some form of potassium compounds, to give a product with a total of 56 fertilizer units. The commercial N-P-K formulas are generally in whole numbers. The chemical nature of the primary element sources are identified elsewhere on the label by the manufacturer. None of the N-P-K formulas totals 100% plant nutrients because the formulas indicate only the nutrient portions of the primary-element compounds and do not account for any other materials present.

Other unidentified materials present in the fertilizer products include free water and water content of the hydrated primary components, uncredited cation and anion portions of salts (sodium, silica, chloride, sulfate, and so on), impurities remaining in the chemical ingredients as processed, and deliberately added inert materials (called ballast), to make whole-number formula compositions. Low-analysis formulas contain substantial proportions of ballast. Formulations prepared to correct specific soil deficiencies may include secondary or micronutrient fertilizer elements. The other than primary elements are identified in the fourth and successive positions of the formula. Formula 16-20-0-14S contains the equivalent of 16 wt % N, 20 wt % P_2O_5, 0 wt % K_2O, and 14 wt % S (elemental sulfur).

Form of fertilizers. The fertilizer elements rarely occur in nature in the elemental chemical form, and most would not be useful for plant nutrition if supplied in that form. Growing plants can assimilate only fertilizer elements in the combined state of inorganic compounds that are amenable to osmotic absorption. Many modern fertilizer materials consist of compounds that are immediately usable by the crops to which they are applied. Others are quickly converted within the soil to forms that can be assimilated. Some fertilizer chemicals are specifically designed to dissolve slowly or to delay reaction within the soil and therefore prolong the release of easily absorbed compounds to provide sustained feeding over the growth cycle of the plants. Fertilizers are derived from many different types of raw materials. The earlier forms of fertilizers deliberately applied to soils were readily available waste products from the activities of other enterprises. Primary, secondary, and trace elements are now compounded into finished fertilizer materials as major products of the chemical process industry. Large amounts of fertilizer chemicals are also generated as by-products in the production of other chemicals.

Production, marketing, and use. The fertilizer industry of the United States is organized into separate segments for producing and marketing nitrogen, phosphate, or potash intermediates and products. The relatively few plants for production of large quantities of these basic materials are usually located near the mining operations. In the case of nitrogen fertilizer products, the large production complexes are located near natural gas sources, which may be wells or major natural gas pipelines. There are overlaps in the business, and some producers make a full line of products, including basic grades of bulk granular N-P and N-P-K materials. High-analysis products are made to minimize freight cost per unit of plant nutrient. Bulk products are transported to market by ship, barge, rail, and truck. Substantial quantities of anhydrous liquid ammonia also move by pipeline from plants in natural gas–producing areas to major fertilizer markets in the north-central and midwestern states. One innovation has been the construction of large refrigerated atmospheric-pressure storage tanks, barges, and oceangoing ships for transport of cold liquid ammonia. Nitrogen solutions are also produced in bulk and shipped by pipeline as well as by conventional carriers. The fertilizer needs of farmers are serviced by blending plants located near their markets. Services by some distributors include soil analysis and rental of application equipment.

UAN solutions. Aqueous solutions of urea, ammonia, and ammonium nitrate (UAN solutions) are used directly by the farmers as well as in the preparation of granular N-P-K products by mixing with other materials, such as normal superphosphate and triple superphosphate. UAN solutions are also spread directly by field application or used to prepare complete N-P-K fertilizer solutions or suspensions. Solution injection into irrigation water is gaining in popularity. Solutions completely free of solids and materials that tend to precipitate in hard water are important to prevent equipment plugging in trickle irrigation systems. Clear solutions are also important in pivotal irrigation systems for circular, quarter-section plots. Low volatility of the fertilizer components of the solutions assures minimum evaporation and most economic use.

Suspension fertilizers. These fertilizers consist of aqueous slurries of fine crystals in saturated solutions that are stabilized by small amounts of gelling materials, such as attapulgite clay. Many grades of high-analysis N-P-K suspensions, including secondary nutrients and micronutrients, are produced by blending plants. Suspension fertilizers can be produced in a simple agitated tank with very low expenditure of energy. The suspensions can be maintained in uniform composition during spreading on the fields, and give better dispersion than granular material.

Nitrogen. Only a few species of bacteria and algae are alone able to directly utilize nitrogen from its stable molecular form in the atmosphere. Many members of the Leguminosae family of plants maintain a symbiotic relationship with specific members of the bacterial genus *Rhizobium*; this allows fixation of atmospheric nitrogen in nodules on the roots of the plants. The rhizobia are capable of prolonged independent existence in the soil, but fixation takes place only when symbiosis is established. Under favorable conditions this fixation can annually provide up to 100 lb of nitrogen per acre (112 kg/ha). Though virgin soils accumulate substantial amounts of nitrogen from the symbiosis process and return of decaying vegetation over long periods of time, commercial cropping rapidly depletes available nitrogen in the soil.

It is probable that on average no more than 50% of the applied chemical fertilizer nitrogen contributes to plant growth. Several factors are responsible for this inefficiency. Nitrates are the most readily absorbed and usable form of nitrogen, but the nitrates are all very water-soluble compounds and the nitrate ions are not retained by reaction in the soils. The portions of nitrate not immediately absorbed by plant roots can also be lost by bacterial conversion to diatomic nitrogen gas (denitrification) under anaerobic conditions, with loss accelerated in warm environment and pH greater than 7. Liquid ammonia immediately vaporizes in the soil and may be lost to the atmosphere in significant proportion by expansion of the gas if it is not applied deep enough to be captured by soil colloids and water. All alkaline soils contain mineral carbonates and bicarbonates that react to form unstable ammonium carbonates. Ammonium ions held on colloids are converted to nitrate by bacterial nitrification and then become subject to the same losses as nitrate applied directly. The presence of urease in practically all soils assures rapid hydrolysis of uncoated urea, and similar organic compounds, with the development of local alkalinity, favoring ammonium carbonate formation. Excessive water from rain or irrigation after fertilization can physically remove soluble nitrogen compounds in the runoff. Significant deficiency in any of the growth requirements, such as sunlight, soil moisture, and other vital fertilizer elements, will reduce the effectiveness of applied nitrogen fertilizer and thus result in loss of nitrogen benefit. Proper attention to the balance of all soil requirements gives the best assurance of optimum response from any fertilizer component. *See* NITROGEN FIXATION.

Ammonia. Anhydrous ammonia is the richest source of nitrogen (82 wt % N) of any of the fertilizer materials available. Ammonia is the basic component of aqueous solution fertilizers because of the increased solubility that the presence of ammonia imparts to other important solution chemicals, such as ammonium nitrate and urea. Ammonia is the most water-soluble of the common gases. A saturated aqueous solution at room temperature contains about 30 wt % ammonia in the form of ammonium hydroxide, NH_4OH, which has a strongly alkaline nature and is similar in chemical behavior to the hydroxides of alkali metals.

Practically all modern nitrogen fertilizer is derived from ammonia synthesized from hydrogen and nitrogen. The high-pressure catalytic process for ammonia synthesis is known as the Haber-Bosch process. The process employs an iron catalyst promoted with small amounts of alkali, magnesium, aluminum, and calcium. The effect of the ammonia synthesis process upon the chemical and fertilizer industries has been profound.

The preparation and purification of synthesis gas for ammonia production is of comparable importance to the synthesis reaction system itself. A major requirement is the economic production of large quantities of hydrogen free of impurities that would cause undesirable side reactions or would be damaging to catalysts. Electrolysis of water gives a high-purity hydrogen product, but it is expensive and delivered at near-atmospheric pressure. The only practical raw materials available for quantity production of hydrogen for the initial operation of ammonia synthesis were coal and coke. Coke was selected because its reaction with steam produced the less complicated raw hydrogen gas stream. Hydrogen production for the process in the United States quickly shifted to sweet natural gas, which was abundant and consisted mainly of methane with few troublesome contaminants. Methane, CH_4, is the richest hydrocarbon feedstock and the simplest to convert to hydrogen by catalytic steam reforming [reaction (1)]. Sulfur is a poison to most of

$$CH_4 + H_2O \rightarrow CO + 3H_2 \qquad (1)$$

the catalysts in the system and must be removed. In the case of natural gas or other hydrocarbons, the sulfur is removed by pretreatment of the feedstock. The preparation of ammonia synthesis gas requires the addition of nitrogen to the net hydrogen stream.

An alternative scheme, used to a lesser extent, is the injection of pure nitrogen obtained from cryogenic separation of air apart from the ammonia synthesis system. The residual carbon monoxide in the crude synthesis gas is then converted to carbon dioxide to the maximum extent possible with two stages of water-gas shift conversion. Carbon dioxide is selectively absorbed from the cooled gas by any one of a number of commercially available acid-gas absorption solutions. The rich solution is stripped of its CO_2 content and recycled to the absorber. Residual traces of carbon oxides in the gas from the absorber are then catalytically converted to methane before compression and delivery of the synthesis gas to the ammonia synthesis system.

The need for fertilizer production in densely populated areas, such as the Netherlands and Japan and in developing countries with rapidly growing populations, but without natural gas supplies, stimulated research for reforming catalysts that could accommodate heavier hydrocarbons. Naphtha, a liquid petroleum product that could be easily transported to the plant sites, was plentiful because of lesser demand for gasoline in those parts of the world and thus more economically available than in the United States. Several ammonia plants were built in the 1960s to use naphtha as feedstock for hydrogen production.

Increasing dependence of the industrial nations upon the oil-producing nations was forcefully brought to attention during the oil embargo of 1973. The sharply increasing price of petroleum and natural gas after that time dictated an intensive search for ways to reduce all energy use, including requirements for the production of ammonia and other fertilizer materials. Special attention has been given to heat recovery and reactor design. Renewed interest in catalyst development brought forth new manufacturing techniques promising significant improvements in properties that allow reduction in ammonia plant operating cost. Several techniques were devised to recover and recycle hydrogen and ammonia from the purge gas that was previously lost or burned as fuel. Pressure swing absorption, permeable-membrane diffusion, and cryogenic separation are examples of unit operations that are being used to advantage to recover products and reactants from purge streams. *See* AMMONIA.

Nitric acid. Nitric acid, HNO_3, is the second-most important industrial mineral acid. Nitric acid is not used by itself as a fertilizer, though the pure liquid contains 22 wt % N. The greatest portion of commercial nitric acid in the United States is used to produce fertilizer-grade ammonium nitrate, which has the N-P-K formula 35-0-0.

Many large-scale attempts have been made to achieve direct oxidation of nitrogen in air. The fast high-temperature oxidation of nitrogen, whether it be the diatomic gas in air or in the ammonia molecule, results in almost exclusive formation of nitrogen oxide, NO. Further reactions to higher oxidation states are favored by much lower temperature. The air oxidation of NO to NO_2 is carried out simultaneously with the disproportionation reactions (2).

$$\begin{gathered} 2NO + O_2 \rightarrow 2NO_2 \\ 3NO_2 + H_2O \rightarrow 2HNO_3 + NO \end{gathered} \qquad (2)$$

This complex pair of reactions is highly exothermic and requires a multitray absorption column, with each tray cooled separately for countercurrent absorption of the nitric acid, HNO_3, in water. Nitric acid forms an azeotrope with water at about 68 wt % HNO_3. Several dehydrating techniques have been developed to concentrate nitric acid above 68%. Most fertilizer processes do not require higher than 60% concentration, and only a small portion of the product is concentrated for other markets. *See* NITRIC ACID.

Ammonium nitrate. Ammonium nitrate, AN, is a white, water-soluble, inorganic salt that contains 35 wt % N. As a fertilizer, ammonium nitrate is dissolved readily by moisture in the soil and diffuses outward from its placement. The nitrate portion of this salt is immediately available to plants through the root system, and the ammonium portion is retained by the soil colloids, where it is oxidized gradually to the nitrate form with the aid of bacterial organisms in the soil.

Ammonium nitrate is made by the simple acid-base reaction of nitric acid and ammonia in aqueous solution, with the heat release of 26,000 calories (109 kilojoules), per gram mole. The heat of reaction is usually used to vaporize the ammonia fed to the process and to evaporate water to concentrate the salt solution produced. After the conclusion of World War II, intensive research brought forth several process variations for the production of ammonium nitrate. The Stengel process produces a low-moisture-content melt in a packed reactor, and the melt which solidifies in a thin ribbon on a water-cooled stainless steel belt was flaked and screened to uniform particle size. Crystalline product is made to a controlled crystal size in a recirculation neutralizer-crystallizer when virgin acid and anhydrous ammonia are used. The crystals are centrifuged from the mother liquor and dried. The neutralization reaction has been carried out under various pressure levels from vacuum to as much as 5 atm (32 kilopascals). Graining is the process of kneading a congealing melt to form round granules which are then cooled, screened, coated, and packaged. Dilute ammonia gas streams are scrubbed in direct contact with acid in towers, and the resultant salt solution is concentrated in an evaporator to the desired melt composition. Solutions concentrated to greater than 99.5 wt % form a melt from which dense prills can be made. Prills produced for fertilizer are made as dense as possible to provide strength in handling and resistance to moisture intrusion. Prilling is the process of forming uniform droplets of a molten material at the top of a tower and cooling the droplets to solid spherical beads as they fall through a rising column of air. A common means of forming the droplets is to pass the molten material through a shower-head type of distributor at the top of the tower. This process makes a very uniform product of size set by the design of the system. Ammonium nitrate can also be granulated in any of several types of granulation devices to particle size as desired to match other materials for bulk blending. Another desirable feature of granules is the greater crushing strength compared to that of prills.

Ammonium nitrate is also a component part of other nitrogen fertilizers, such as ammonium sulfate-nitrate. ASN (26% N), calcium-ammonium nitrate, CAN (28% N), and ammoniated nitrophosphate. A troublesome characteristic of ammonium nitrate is the several crystal phases it exhibits with change in temperature at 257, 183, 90, and −0.4°F (125, 84, 32, and −18°C). These different crystal structures have significantly different densities, and the transition from one form to another tends to fracture the solids. Temperature cycling (particularly above and below 90°F or 32°C) can reduce well-shaped flakes, prills, or granules to powder within storage piles or containers. Much research has been done to alleviate this problem.

Between 1950 and 1980 ammonium nitrate was the most important nitrogen fertilizer from the standpoint of world consumption. However, in spite of its continued application in large quantity, the percentage use has been steadily declining relative to the use of urea as a source of nitrogen. A large amount of ammonium nitrate is used to make nitrogen solutions for direct application, for the preparation of slurries, and for increasing the nitrogen content of solid fertilizers in granulation operations. The presence of ammonia increases the solubilities of ammonium nitrate and urea in aqueous solution, and a number of high-analysis nitrogen solutions are made by mixing ammonium nitrate, ammonia, and urea. The solutions are simple to prepare and easy to apply to farmlands. *See* PRILLING.

Sodium nitrate. Sodium nitrate is a fertilizer that provides 16% N, and has been particularly favored for cotton and beets. While natural deposits (caliche) of sodium and potassium nitrates have been reported in widely scattered parts of the world, the only deposits of commercial significance are in the northern desert of Chile. Caliche is a complex mixture of natural double salts of sodium, potassium, calcium, and magnesium. Nitrates and sulfates predominate, but recoverable portions of chlorides, iodides, and borates are also present.

An early high-temperature recovery method, known as the Shanks process, was introduced in Chile late in the nineteenth century. Another process came into operation in 1922. This was the Guggenheim process, which was designed to treat low-grade ores with efficient nitrate extraction at moderate temperature and crystallization with mechanical refrigeration. This process soon demonstrated the ability to economically treat ore with nitrate concentration as low as 7%. Large-scale operations were initiated in the late 1920s and early 1930s which greatly extended the life of Chile's deposits. The recovery of iodine and other mineral values in the caliche also enhanced the economics and helped to extend the operational life.

Following World War I, considerable effort was exerted to devise methods for producing synthetic sodium nitrate to guard against industry crisis by interruption of supply of the natural product from Chile. The five different process schemes that were commercialized used either nitric acid directly or oxides of nitrogen produced by the catalytic oxidation of ammonia. Sodium was supplied in the form of soda ash, common salt, or caustic soda. In one of the processes the reaction of nitric acid with salt (sodium chloride) also produced chlorine without generating by-product sodium hydroxide, as in the case with electrolytic production of chlorine. All the synthetic processes yielded aqueous solutions from which the sodium nitrate was recovered by crystallization. Production peaked in the 1930s and continued brisk through the 1940s. The fertilizer use of sodium nitrate was then gradually replaced by ammonium nitrate and other nitrogen materials. Some sodium nitrate is being used, but it is unlikely that it will again enjoy its previous favored position in the fertilizer market.

Other nitrogen salts. Ammonium sulfate, $(NH_4)_2SO_4$, is a water-soluble, crystalline material containing 21% N and 24% S. Ammonium sulfate was first made in substantial amounts by the absorption of the ammonia content of coke-oven gas in scrubbers with aqueous solution of sulfuric acid and subsequent crystallization of the salt from the concentrated scrubber liquor. It is also made by reacting calcium sulfate with ammonium carbonate, filtering out the precipitated calcium carbonate, and crystallizing ammonium sulfate from the mother liquor. Large quantities of ammonium sulfate have been produced by the reaction of fresh ammonia with spent sulfuric acid from chemical operations and oil refinery catalytic alkylation processes. A substantial amount of ammonium sulfate has been made by the reaction of virgin sulfuric acid and ammonia. Ammonium sulfate produced as by-product in the manufacture of caprolactam and other large-tonnage chemicals has supplemented the world supply of fertilizer nitrogen. The intensive interest during the 1970s in environmental protection from industrial sulfur dioxide emissions led to development of several processes for recovery of the sulfur values, which include production of ammonium sulfate. Most of this pollution control by-product also finds its outlet in the fertilizer market.

Ammonium chloride or sal ammoniac, NH_4Cl, is a water-soluble, crystalline salt with 25 wt% N. Ammonium chloride is more corrosive than other ammonium fertilizers, and it also tends to increase the residual chloride content of soils. In spite of these characteristics, it has been a popular fertilizer material for rice and some field crops. A rice fungus that reduces sulfates to toxic sulfides under anaerobic conditions does not affect chlorides, and thus ammonium chloride is preferred over ammonium sulfate in such applications.

The salt can be easily made by reacting ammonia with hydrochloric acid and crystallizing the product. However, the major manufacturing method for ammonium chloride production is the double-salt process, which is a modification of the Solvay ammonia-soda process for making sodium carbonate. A sodium chloride solution is ammoniated and blown with carbon dioxide in a carbonating tower until the ammonia is carbonated to the bicarbonate, whereupon it reacts as formed with sodium chloride to give sodium bicarbonate and ammonium chloride. The Solvay process reaches equilibrium at about 75% conversion in the salt exchange, sodium bicarbonate is recovered by filtration, and the ammonia is recovered by the addition of calcium hydroxide, leaving a solution of calcium chloride. The ammonium chloride in the double-salt process regenerated after separation of the sodium bicarbonate is salted out by ammoniating the mother liquor, cooling, and adding fresh sodium chloride. The ammonium chloride is removed by centrifuge, and the cycle is repeated. The sodium bicarbonate is then calcined to sodium carbonate, and the carbon dioxide from calcination is used to carbonate the ammoniated salt solution in the tower.

Calcium cyanamide. Calcium cyanamide, $CaCN_2$, is an alkaline fertilizer material with fixed nitrogen. It is black and dark gray because of the residual unreacted carbon remaining from the reduction step in its production. The fertilizer nutrient content is 35% N and 70% CaO. Decomposition after application to the soil converts calcium cyanamide to calcium carbonate and ammonia according to reaction (3). It has additional agricultural value as a herbicide and as a defoliant for mature cotton plants to aid machine harvesting of the fiber. *See* DEFOLIANT AND DESICCANT; HERBICIDE.

$$CaCN_2 + 3H_2O \rightarrow CaCO_3 + 2NH_3 \qquad (3)$$

Three manufacturing steps are involved in the production of calcium cyanamide. The basic raw materials are limestone and coke. Limestone, $CaCO_3$, is first calcined to quicklime, CaO. Calcium carbide, CaC_2, is then produced in an electric arc furnace (3600–4000°F or 2000–2200°C) by the reduction of quicklime with coke. This reaction is highly endothermic. The calcium cyanamide is then made by passing nitrogen gas, N_2, directly over heated calcium carbide in a furnace. This high-temperature reaction (1650–1830°F or 900–1000°C) is sufficiently exothermic to be self-sustaining, once it is iniated. Oxygen-free nitrogen is supplied from an air separation plant. The major use of calcium cyanamide since its discovery in 1898 has been agricultural. Its cost as a fertilizer is high in today's market relative to other fixed-nitrogen materials. Further agricultural use of calcium cyanamide is therefore likely to be limited to special services that can justify its high cost.

Urea. Fertilizer-grade urea is a water-soluble, synthetic, organic solid that contains 46% N. It has the highest nitrogen nutrient value of any of the solid fertilizers. Urea is less deliquescent than ammonium nitrate, but the critical relative humidity is 72% at 86°F (30°C), and so urea must be stored in a controlled low-relative-humidity environment to remain free-flowing. The many good characteristics of urea as a nitrogen fertilizer are responsible for its rapid worldwide acceptance. Urea is a stable material with a convenient melting point (270.9°F or 132.7°C) that allows it to be easily prilled or granulated for bulk blending with other solid fertilizer materials. Urea has high solubility in water and can be used to prepare high-analysis solutions. It forms an adduct with phosphates which allows the production of homogeneous granular solids. Low-cost by-product carbon dioxide from ammonia production related to making urea helps keep the cost competitive with other nitrogen materials. These two production units are commonly associated in a single fertilizer complex facility.

The process for urea synthesis in commercial use today is the pressure synthesis from ammonium carbamate. Ammonia and carbon dioxide are reacted under pressure of 170–400 atm (1.2–2.8 megapascals), with melt temperature maintained in the range of 347–410°F or 175–210°C. The selection of operating conditions is a compromise among several interactive factors. Yield is favored by high temperature, high pressure, and minimum excess ammonia. Unfortunately, these same conditions are also the worst conditions from the standpoint of equipment corrosion, and much attention has been given to this problem. The partial dehydration of ammonium carbamate formed in the reactor produces urea and water. The mixture let down in pressure from the reactor is stripped of ammonia and carbon dioxide, and the aqueous solution of urea is concentrated for direct granulation or prilling. The simplest and least expensive design for a plant is once-through operation. Then the off-gas containing excess ammonia from the reactor and the carbamate decomposer are used to make other products in associated process units. Plants with various degrees of gas and liquid recycle have also been constructed. Most plants today are designed for total recycle with maximum possible heat recovery. Some licensors offer plant designs having several levels of pressure letdown with vapor absorption and solution recycle arranged to minimize process energy consumption.

The hot urea melt is time-temperature-sensitive with respect to decomposition by condensation of 2 moles of urea to biuret with the loss of 1 mole of ammonia. Many crops show a toxic response to biuret, and fertilizer-grade urea has strict limits on the allowable biuret content. Most urea is concentrated by evaporation of water with the minimum possible melt residence time before entering the prill tower distributors to keep the biuret content below 1 wt %. Especially low-biuret urea can be made by the more expensive concentration with low-temperature crystallization from the aqueous solution, and prilling or granulating the melted crystals. Granulation avoids the atmospheric emission problem of prilling and has the advantage of adjusting particle size to match other solids for bulk blending. The matching of particle size is important to prevent resegregation of the different materials in the blend after packaging or in the bulk blend transported to the field. The granulated product also has much higher crushing strength than the prills. Sulfur-coated urea is made by spray-coating granular urea with molten sulfur in a rotating drum and seal-coating the product with a wax or low-molecular-weight polyethylene. The smaller-diameter urea prills are unsuitable for coating because the high surface-to-volume ratio would require too much sulfur and result in product too low in nitrogen content. Treating urea in this manner has several advantages. In addition to providing controlled release of nitrogen and sulfur, the coated material has excellent storage and handling characteristics that allow its use in bulk blending in humid climates. It may also be blended with single or triple superphosphates, which are incompatible with uncoated urea. *See* UREA.

Ureaform. Ureaform contains at least 35% N in a largely insoluble form that slowly and continuously becomes available in the soil. It is one example of the several slow-release fertilizer materials on the market. In the late 1930s an aqueous ammoniating solution of urea and formaldehyde was used to ammoniate superphosphate in the manufacture of mixed fertilizer. This operation produced a water-insoluble reaction product of urea and formaldehyde called ureaform. The advantage of ureaform as a fertilizer is its relative insolubility in cold water. This slows the nitrification process in the soil and thus provides an extended release of usable nitrogen to the plants. The condensation of urea and formaldehyde with the loss of water is catalyzed by acid in aqueous solution. Technology developed rapidly, and it is now made commercially by either the dilute solution process or the concentrated solution process. In either process the granulated and dried product is an odorless, white solid containing about 38% N.

Ureaform is not a single compound, but a mixture of methylene urea polymers that are capable of continued polymerization at reduced pH and elevated temperature. Ureaform is not hygroscopic. The conditions of manufacture strongly influence the quality of the product as determined by the activity index, AI, which is the percentage ratio of polymer soluble in hot water (212°F or 100°C) to the polymer insoluble in cold water (77°F or 25°C). As the degree of polymerization increases beyond the desired level, the hot-water solubility of the ureaform decreases. The best-quality ureaform has a high percentage of cold-water-insoluble material and a high AI value. This property makes ureaform an ideal direct-application nitrogen fertilizer for turf grass, ornamental shrubs, and greenhouse plants. A single annual application can satisfy the nitrogen requirements without the danger of burning the crop, as can be the case when more active nitrogen fertilizers are applied in excess or with too little water. Ureaform is also employed in batch and continuous processes to make dry-mix products and homogeneous granular N-P-K fertilizers. Care must be taken during these operations to avoid prolonged exposure to high temperature or to pH outside the range of 5.4–6.2. Urea is the preferred nitrogen supplement. Pulverized dolomitic limestone and hydrated lime are commonly used to compensate for the acidity of granulation ingredients such as superphosphate. Storage temperature of these N-P-K products must be maintained below 100°F (38°C) to avoid degradation of the ureaform.

Phosphorus. Phosphorus is a constituent of lecithin and other lipoids in plants. The average phosphorus, P_2O_5, content of whole plants is about one-third the nitrogen content. Highest concentrations are found in the germ and bran of grain crops. Phosphorus is the eleventh most abundant element in the Earth's crust. The elemental form of phosphorus does not exist in nature. Its minerals are widely distributed in the Earth's crust in moderate concentrations. Many deposits in various parts of the world are used in commercial production of phosphate rock. Numerous forms of phosphate materials, including untreated phosphate rock, have been used for fertilizer. Large amounts of phosphates are needed just to replace the quantity carried to the sea by rivers in the form of land erosion, sewage discharge, and other chemical losses. It is estimated that combined annual phosphate losses to the sea from the United States alone are equivalent to about 60×10^6 tons (54×10^6 metric tons) of phosphate rock.

Phosphate rock. The principal source of phosphate for the fertilizer industry is phosphate rock, which is a generic name for ores of phosphorus-containing minerals. The chief phosphatic component of these ores is fluorapatite, with the general formula $3Ca_3(PO)_2 \cdot CaF_2$. The name phosphorite is given to sedimentary apatites. The phosphate content of the amorphous phosphorites is more easily released by the action of mineral acids than is that of the hard crystalline apatite. The phosphorites are therefore the favored commercial ores. The phosphate content of phosphate rock is frequently given in bone phosphate of lime (BPL) or % $Ca_3(PO_4)_2$, as well as in terms of P_2O_5. A high-grade rock of 75 BPL is equivalent to 34.4 wt % P_2O_5. Large tonnages of ground phosphate rock have been used directly for soil fertilization without further treatment. The availability of the P_2O_5 content of this phosphorite application is dependent upon the pH and mineral content of the soil. Ground phosphate rock is the least costly phosphate fertilizer, but release to the crop is slow under the best conditions. *See* APATITE.

Phosphate rock as mined is usually beneficiated (concentrated) to remove gangue and increase the P_2O_5 content for feed to other processes. Beneficiation includes washing, screening, crushing oversize, treating fines by flotation, and agglomerating the phosphate concentrage.

Superphosphate. To produce superphosphate, concentrated sulfuric acid, water, and ground rock are continuously proportioned into a mixer which intimately blends the reactants. The hot acidulated slurry from the mixer, such as the Kuhlman turbine which emulsifies the ground-rock acid, discharges into a continuous den. The den encloses a slow-moving conveyor, or a rotating plate where the slurry sets up to a spongy mass (consisting mainly of monocalcium phosphate hemihydrate, $Ca(H_2PO_4)_2 \cdot H_2O$, and calcium sulfate hemihydrate, $CaSO_4 \cdot {}^1/_2H_2O$) with the evolution of gas. At the end of travel the plastic superphosphate mass is broken up by a cutter and discharged onto a belt conveyor which delivers the product into storage to cure. The curing process allows complete reaction of the sulfuric acid by penetration of the phosphate rock, which maximizes the availability of the P_2O_5 content. Superphosphate recovered from curing storage is crushed and screened for sale as run-of-pile.

The superphosphate is nearly always dark in color because of residual carbon left in the product from charring of organic material in the rock and the spent acid during the acidulation reaction. Ammonia will react with monocalcium phosphate according to reaction (4). The heat of this reaction and the cementing action of ammonium phosphate formed allow granulation of the mixture with little additional expenditure of energy. Overammoniation will cause a reversion to $Ca_3(PO_4)_2$ and a loss of P_2O_5 availability. The fertilizer value of the ammoniated superphosphate enhanced with the nitrogen content caused a strong demand for this product. The simplicity of the reaction and the effectiveness of the product allowed superphosphate to remain unchallenged in the top position of phosphate fertilizers for more than a hundred years from the time of the first patent, issued in 1842. Superphosphate, sometimes called normal or single superphosphate, continued to represent half of the world production of phosphatic fertilizers as recently as 1962. However, superphosphate fell in popularity because of its low P_2O_5 content and the rising cost of freight. To a large extent, it has been replaced by high-analysis fertilizers such as a triple superphosphate (TSP) and the ammonium phosphates (MAP and DAP).

$$Ca(H_2PO_4)_2 \cdot H_2O + NH_3 \rightarrow CaHPO_4 + NH_4H_2PO_4 + H_2O \quad (4)$$

Phosphoric acid. Phosphoric acid is a major commodity of the fertilizer industry and the chief source of phosphorus in modern commercial fertilizers. It is a strong mineral acid and is not applied directly to the soil as a fertilizer. The merchant grade of orthophosphoric acid, H_3PO_4, contains 54 wt % P_2O_5. *See* SOLVENT EXTRACTION.

1. *Production*. Phosphoric acid is made by either of two general routes: the electric furnace process or acidulation of phosphate rock.

Furnace acid, or white acid, is made by the electric furnace process. Coke and silica are mixed with coarse phosphate rock to provide a porous bed that is melted by heating to a high temperature with electric current passed through the electrodes and carbon blocks forming the bottom of the slag pool in a closed furnace that also serves as a retort. The phosphorus is reduced to the element, which vaporizes in the furnace. Dust is removed from the vapor with electrostatic precipitation. Air is excluded from the system, and the vapor is condensed to liquid white phosphorus by direct contact with water sprays. The liquid phosphorus is collected under a layer of water and pumped to a burner where it is oxidized with air to phosphorus pentoxide, P_2O_5, in a water-cooled, graphite-block, combustion chamber. The white fumes of P_2O_5 are absorbed in water or dilute acid in a spray tower to produce water-white, high-purity orthophosphoric acid. The combustion and hydration can also be done in a single tower. Because other minerals remain with the slag in the furnace, the furnace process can accept lower-grade phosphate rock and can tolerate large amounts of silica and more trivalent metal contaminant than the wet process. Ferrophosphorus is a by-product of this process that can be used for steelmaking. The furnace reduction process is important for production of elemental phosphorus for munitions and some important chemicals. The white acid was used for the production of high-analysis fertilizers when electric power was cheap and plentiful.

Wet process, or green, phosphoric acid is made by the acidulation of phosphate rock with dilute solutions of strong mineral acids, such as sulfuric, hydrochloric, or nitric. Practically all the phosphoric acid used in modern fertilizer production is made by the wet process. Sulfuric acid is most commonly used in this process. There are several variations in design among the proprietary processes, but in general, finely ground phosphate rock is digested with diluted acid in an agitated reactor, or in a series of them, where phosphoric acid is formed along with calcium sulfate, which is precipitated in a manner to yield a filterable dihydrate crystal slurry, according to reaction (5). A slight excess of sulfuric acid

$$3Ca_3(PO_4)_2 \cdot CaF_2 + 10H_2SO_4 + 20H_2O \rightarrow 6H_3PO_4 + 10CaSO_4 \cdot 2H_2O + 2HF \quad (5)$$

is used to maximize phosphorus recovery. The relatively insoluble calcium sulfate dihydrate (gypsum) crystals are filtered and washed with two or three progressive washes of filtrate followed by a water wash. The strongest wash is returned to the digestor with fresh sulfuric acid feed. The washed gypsum is usually discharged to a storage pond in a water slurry, but may be used for ballast in granulation of N-P-K fertilizers or reprocessed for manufacture of cement and wallboard. The filtered phosphoric acid is recovered at 28–30% P_2O_5, which is then usually concentrated in rubber-lined vacuum evaporators to merchant-grade acid. When the phosphate rock supplied to this process is first calcined to burn off the organic matter in it, the acid from the filter often has a rich green color due to the metallic elements in the ore. This coloration gave rise to the name green acid for wet-process phosphoric acid. The organic matter in uncalcined phosphate rock used in the digestion becomes charred by the reaction, and the colloidal carbon particles give the phosphoric acid a dense black color. In either case, the acid produced by this process is supersaturated with respect to a number of complex salts of iron, aluminum, magnesium, potassium, fluorine, and silica. The continuous precipitation of these complex salts has caused problems in storage and shipment of the acid. Much effort has been spent by industry to find a means of preventing the formation of the hard deposits of these materials.

2. *Purification*. Among the many purification processes for wet-process phosphoric acid produced with sulfuric acid, urea phosphate crystallization is the simplest and has the lowest energy demand. The calcination of phosphate rock to remove organic matter before acidulation is an effective means of achieving clear wet-process phosphoric acid. However, this energy-intensive process is expensive and is practiced only with phosphate rock containing high concentration of organics. Several proprietary extraction and ion exchange processes have been developed for the purification of phosphoric acid produced by rock acidulation with hydrochloric acid and with nitric acid. These processes are all expensive with respect to plant investment and operating costs. The limited cases of their application have resulted from unusual circumstances. It is unlikely that any of these processes will be in widespread use.

3. *Uranium recovery*. Phosphate rock contains a small percentage of uranium that can be recovered from wet-process phosphoric acid by organic solvent extraction. Much of the phosphoric acid produced in the United States is now being profitably extracted for uranium recovery. The dramatic increase in energy cost has prompted renewed efforts to improve the economics of phosphoric acid production. One variation in the design of the wet process is crystallization of calcium sulfate hemihydrate, $CaSO_4 \cdot {}^1/_2H_2O$, in the acid digestion, which occurs at higher temperature and greater P_2O_5 concentration than is the case with gypsum formation. Filtration of the hemihydrate requires greater filter surface because the crystals do not drain as rapidly as the gypsum crystals. Other variations include the foam process and the two-step recrystallization of hemihydrate to gypsum. The goal of these variations is the production of phosphoric acid of sufficient strength (40% P_2O_5) for granulation without further concentration of the filtered acid. Most facilities that manufacture phosphoric acid use all or part of the product captively to make granulated N-P-K fertilizers.

4. *Phosphate rock supply*. The highest grades of phosphate rock are preferred for long-distance shipments because freight has become an ever larger fraction of the delivered cost. The large demand for phosphate fertilizers in the past few decades has led to depletion of the highest grades of phosphate rock in currently active mines. Very large reserves of phosphate rock remain in the United States and other producing parts of the world, but the once readily available 75 BPL rock is in short supply and 68 BPL is now commonly shipped long distances. Even more important than P_2O_5 content of the rock to the production of wet-process acid is the contamination of specific minerals, of elements such as iron and aluminum that cause loss in production by troublesome precipitations.

5. *Superphosphoric acid*. Superphosphoric acid is the viscous product formed by further concentration of merchant grade to increase the P_2O_5 content to 70% or more. Wet-process phosphoric acid at 54% P_2O_5 concentration is equivalent to about 75 wt % H_3PO_4, with a viscosity of about 12 centipoise (0.012 pascal-seconds) at 100°F (38°C), and practically all phosphate is in the form of orthophosphoric acid. As the acid solution is further concentrated by the loss of free water, the temperature rises and the phosphoric acid gradually begins to polymerize to the dimer, pyrophosphoric acid ($H_4P_2O_7$), and higher condensed forms by the progressive loss of combined water. The viscosity increases rapidly with the increase of mineral contaminants. The viscosity of furnace-grade superphosphoric acid containing 40% polyphosphate at 100°F (38°C) is about

250 centipoise. The viscosity of wet-process superphosphoric acid also containing 40% polyphosphate at 100°F (38°C) will vary between 2000 and 20,000 (2 and 20 Pa · s) centipoise, depending upon the amount of impurities present. The superphosphoric acid must therefore be stored and handled in heated systems to retain fluidity. The furnace-grade superphosphoric acid is made by using less water in the absorption tower receiving the P_2O_5 from the burner. The temperature range of superphosphoric acid concentration (446°F or 230°C) is well above the acceptable limit for rubber-lined equipment, and so the concentration must be done in expensive alloys or carbon brick equipment. One superphosphoric acid concentrator design holds the acid in a conical pool of a chamber lined with carbon brick while it is heated by submerged combustion or the direct contact with hot combustion gases blown through the pool of acid. The large amount of acid mist generated by this operation is recovered by means of a high-energy venturi scrubber followed by a stationary entrainment separator. In addition to the high P_2O_5 content of superphosphoric acid made from wet-process acid, it has the ability to sequester impurities in the wet-process acid and prevent their troublesome precipitation.

Triple superphosphate. TSP, also called concentrated superphosphate, is made by the action of phosphoric acid on ground phosphate rock to give a product that is predominantly water-soluble monocalcium phosphate monohydrate. Unlike single superphosphate, TSP is not diluted with calcium sulfate, so that the P_2O_5 concentration is generally 40–49%, depending upon the quality of the rock used. The presence of other minerals in the rock leads to the formation of some citrate-soluble phosphates, such as iron and aluminum phosphates. These and other trace elements are of benefit to granulation of the TSP and contribute to micronutrient fertilizer value. Citrate solubility of phosphates is acceptable as P_2O_5 availability in the United States, but the legal requirement of water solubility is much more important in Europe and the United Kingdom.

Acidulation of phosphate rock with phosphoric acid releases silicon tetrafluoride and hydrofluoric acid, as does the acidulation with sulfuric acid, but the amount of fluoride released is dependent upon the phosphoric acid concentration, which also influences the temperature of the reaction with the phosphate rock. The speed of the reaction and the setting time of the reaction mass are also determined by the temperature and acid concentration.

Several processes have been developed which include liquid and solid reaction types with batch or continuous operation to make powdered or granular product. The equipment and reactions are similar to those for single superphosphate, but the mixing periods are shorter and the reaction mass solidifies more quickly in the production of triple superphosphate. The initial moisture content is greater and the curing time is longer. Most TSP is manufactured by continuous granulation of the reaction products, without the intermediate curing in a pile. The heat from the dryer in the granulation circuit speeds up the curing process, and the dry granular material is conveyed to storage. Shipment of TSP is seasonal in many parts of the world, and large bulk storage must be provided.

Ammonium phosphate. Wet-process monoammonium phosphate, $NH_4H_2PO_4$ (11-48-0), and diammonium phosphate, $(NH_4)_2HPO_4$ (18-46-0), have become popular fertilizers because of their high fertilizer analyses and good storage and handling characteristics. Both materials are commonly made in granulation circuits by reacting anhydrous ammonia with wet-process phosphoric acid. The high degree of ammoniation requires carefully designed scrubbing systems to recover and recycle excess ammonia carried in the air streams from the granulator and the dryer along with dusts generated in the solids-handling equipment. The fresh incoming acid is usually used to scrub the ammonia vapors and collect the dust. This partially ammoniated acid also containing dissolved ammonium phosphate dusts is then further ammoniated in a preneutralizer. The preneutralized hot slurry is distributed onto a bed of semiwet granular solids where more ammonia is introduced within the bed to react with phosphoric acid on the surface of the granules. The moist, blended granules are dried in a cocurrent rotary dryer and screened on double deck screens. The oversize particles are crushed and returned to the granulator along with the fines. The intermediate-size particles are withdrawn as required for production, cooled, and conveyed to storage, while the excess is returned to the granulator.

Many different types of reactor circuits and scrubbing systems have been designed to achieve these purposes. Two granulation devices most used are the ammoniating drum granulator and the pug mill granulator, or blunger. The drum granulator produces the harder and more uniformly round granules. Either type of circuit can be used to make a full range of N-P-K fertilizers by adding potassium compounds to the granulator feed. Consumption of the ammonium phosphates has steadily increased since the mid-1950s till these materials now constitute over 60% of the world usage of P_2O_5.

The SAI process makes a stable powder form of monoammonium phosphate (MAP) containing 6–12% moisture that has good storage and handling properties similar to superphosphate. The MAP is made with a slurry of wet-process phosphoric acid ammoniated to 1.35 N:P ratio, which is then reduced to 1.0 N:P ratio by the addition of more acid. The great difference in solubility of saturated ammonium phosphates causes the mixture to congeal with minimum solubility at the composition of MAP by the addition of phosphoric acid to the very soluble mixture of DAP and MAP. Mineral impurities in the wet-process acid influence the crystal growth of MAP in the process. The congealed MAP is broken up and dried at low temperature (257°F or 125°C) to no less than 6% moisture, which retains the good granulating characteristics of the material. Development of the pipe-cross reactor made possible the direct contact of ammonia vapor with phosphoric acid or sulfuric acid at high temperature. This high reaction temperature allows ammonium phosphate production from orthophosphoric acid without the need for a preneutralizer. Much of the water in the phosphoric acid is evaporated directly in this reactor with reaction heat, which reduces the recycle rate and the energy requirement in the granulation circuit. The low-moisture-content melt (less than 2%) fed to the granulator from the pipe-cross reactor also eliminates the need of a dryer for some N-P-K grades. The reduction of materials handling within the process unit also reduces air-pollution problems. The favorable economics of this operation has caused many producers to retrofit granulation plants with the pipe-cross reactor to manufacture high-analysis granular N-P and N-P-K products.

Potassium. Potassium has several important functions in plant metabolism. It appears to play an essential role in the water economy of plants through the action of its soluble compounds on osmotic pressure. Potassium is involved in oxidative reactions as a carrier for iron required by important enzymes that convert sugar to starch and amino acids to protein. Sugarbeets, potatoes, and other plants that produce large quantities of carbohydrates use large amounts of potassium.

Complex potassium minerals are widely distributed, and most are insoluble. However, some extensive mineral deposits of single and double salts of chlorides and sulfates are soluble and provide the major sources of potassium supply. Most of the potassium supplied to the fertilizer industry is in the form of the commercial-grade crystalline chloride salt (60% K_2O). Potassium sulfate (50% K_2O) is supplied for some crops that require potash fertilization, such as tobacco, and are sensitive to chloride. The potassium salts are usually mixed with other fertilizer materials in dry-blended formulas or in homogeneous granules. Solution and suspension fertilizers containing potassium compounds gained in popularity because of lower cost resulting from processes that require less energy than those involving granulation or drying and blending of solids.

Potassium salts are also present in minor concentration in the caliche deposits of Chile. Development of the Guggenheim process made possible the economic recovery of potassium nitrate, KNO_3, along with other minor minerals in conjunction with production of sodium nitrate. Potassium nitrate is an expensive, but totally soluble high-analysis fertilizer that combines both N and K_2O in a single material containing 60 units of plant nutrients. It is favored for special formulations that are applied in solution form to some chloride-sensitive crops. *See* POTASSIUM.

Sulfur. Sulfur is widely distributed in plant tissue. It is a component of the amino acids methionine and cysteine, and is also found in some hormones and in oil glycosides that impart odor and flavor to mustard, garlic, and onions.

Large deposits of elemental sulfur trapped geologically in deep salt domes along the Gulf coast of the United States and Mexico are economically mined with superheated steam by the Frasch process. Sulfur is recovered in variable quantities in the form of hydrogen sulfide from natural gas, once burned in large quantity in flares, that is now recovered as elemental sulfur by the Claus process. Environmental regulations have brought about the recovery of sulfur dioxide from the metallurgical processing of sulfide ores of iron, copper, lead, zinc, mercury, and antimony. It is usually most economic to convert the recovered sulfur dioxide directly to sulfuric acid, provided that the acid can be used close to the site of production. The recovery of sulfur dioxide from the combustion of oil and coal is an environmental necessity that is a source of sulfur and sulfuric acid for the fertilizer industry. Enormous deposits of high-quality gypsum (calcium sulfate dihydrate) are located close to the surface and available for recovery of sulfur values in the event that the more easily converted forms of sulfur should become insufficient for world needs. *See* GYPSUM.

Most soils contain substantial amounts of sulfates; and fertilizers such as superphosphate, ammonium sulfate, and ammonium sulfophosphate, commonly used in the past separately or in mixtures, had ample sulfur replacement value to maintain growth requirements. These and other fertilizer materials were generally credited for their primary fertilizer elements, and the presence of other nutrient elements, such as sulfur and calcium, were ignored. The trend to high N-P-K analysis fertilizers to offset the rising cost of freight led to evidence of sulfur deficiency in some heavy-sulfur-consuming crops. Sulfur deficiency manifests itself by marked chlorosis and retardation of plant growth. The extensive switch to triple superphosphate and ammonium phosphates from normal superphosphate has practically eliminated the generous amount of sulfate that was formerly provided without charge or recognition and applied to crops. Continued and expanded evidence of sulfur deficiency has brought about the focusing of greater attention to the secondary elements in commercial fertilization programs. From a tonnage standpoint, sulfur is even more important to the fertilizer industry in the form of sulfuric acid as a means of providing phosphorus availability from phosphate rock. This is by far the biggest worldwide use of sulfur. *See* SULFUR; SULFURIC ACID.

Calcium. It is believed that calcium is utilized in continuous cell division and is involved in nitrogen metabolism because the absorption and assimilation of nitrates in plants suffer retardation when calcium is deficient. Calcium is not a mobile element in plants in the manner of phosphorus. Besides the fertilizer function of calcium, it is applied to the soil to correct excessive acidity and improve the physical structure of the soil. The application of lime to an acid soil in an amount to change the pH from 6.5 to 7.5 creates a condition favorable to the normal activities of beneficial soil microbes. Calcium is leached out of soils in the form of calcium biocarbonate. The trend to produce very high-analysis fertilizers has resulted in severe reductions in calcium that was previously applied along with N-P-K formulas prepared from normal superphosphate. Though calcium is a widely distributed element in nature, the practice of eliminating its inclusion in fertilizers normally applied to the fields may in the future result in deficiency that is not evident at present. Calcium deficiency is less likely to appear in Europe where the acidulation of phosphate rock with nitric acid is still practiced and nitrogen is frequently applied in the form of calcium nitrate and calcium ammonium nitrate and some calcium cyanamide. Phosphate fertilizer in the form of basic slag from steel production contains about 17 wt % P_2O_5 as tetracalcium phosphate and other compounds of calcium. Basic slag is a common fertilizer item in Europe but of little importance in the United States. Phosphate rock is a rich source of calcium, and it is also available in huge natural deposits of limestone and gypsum. *See* CALCIUM.

Magnesium. Magnesium is the central atom in the structure of chlorophyll, which is vital to photosynthesis; most of the magnesium found in plant tissues is concentrated in the green parts, which contain maximum chlorophyll. Magnesium is credited with translocation of phosphorus in the plant and participation in the activity of enzymes. The disruption of plant processes due to a deficiency of magnesium is manifested by chlorosis, or yellowing of the green tissues. Magnesium is the eighth most abundant element in the earth's crust and is widely distributed in the soil in the form of carbonates and silicates. Magnesium minerals are more soluble than their calcium counterparts, and leaching losses occasionally cause a soil deficiency. Magnesium supplements are often supplied by liming with dolomite, $CaMg(CO_3)_2$. Other minerals commonly used as magnesium fertilizer are langbeinite, $K_2SO_4 \cdot 2MgSO_4$, and magnesium sulfate. *See* CHLOROPHYLL; MAGNESIUM; PHOTOSYNTHESIS.

Micronutrients. The micronutrient elements essential to plant growth are usually present in the soil in concentrations adequate for the small quantitative demands of the crops. Natural distribution of the elements is, however, nonuniform, and some shortages do occur. Also, long-term usage of the land can deplete the natural supply of these elements. Moreover, it is possible that applications of pesticides and fertilizer materials to supply primary and secondary elements may react in a way to reduce the availability of one or more of the micronutrients. Replacement of these fertilizer elements required in trace amounts poses a problem because of the narrow concentration margin between deficiency and toxicity. It is necessary that the desirable compound of the micronutrient be diluted in the applied material to avoid local high concentrations that would occur in the soil in the immediate vicinity of a concentrated granule of the micronutrient. This is achieved by adding a small percentage of the compound in the manufacture of granulated N-P-K formulas, liquids, or slurry mixtures. These dilute forms can then be distributed over the soil to assure uniformity of application. The distribution of a dry blended mixture of solids of the same overall concentrations cannot achieve the necessary degree of dispersion that is possible with the diluted materials.

Boron. Boron is an essential micronutrient that is required in very low concentration. The role played by boron in the life process of plants is unknown, but specific deficiency symptoms traced to shortages of boron include leaf-roll of potato, browning of cauliflower, and brown heart of turnip and sugarbeets. The desirable soil concentration ranges very considerably from crop to crop. Tomatoes do best with boron concentration below 1.0 ppm, while asparagus development is favored with boron concentration as high as 15 ppm. Boron-deficient soils can be adjusted by applications of borax (sodium tetraborate), boric acid, or boron frits in the order of 2.7–16 lb/acre (0.5–3.0 kg/hectare). Faster corrections can be made with foliar applications, but the range of concentration between optimal feed and toxicant is narrow. *See* BORON.

Cobalt. Cobalt concentrations in the range of 0.1–0.3 ppm are important to healthy plant growth, but the function of cobalt in plant metabolism is obscure. Fertilizer forms of cobalt are the soluble salts of sulfate, chloride, or nitrate, which may be applied admixed in aqueous solutions slurries, or homogeneous granules of N-P-K products. *See* COBALT.

Copper. Copper is an important nutrient that must be present in very low concentration. Copper has an essential, though not fully understood, role of participation in oxidizing-reducing enzymes, such as tyrosinase and ascorbic acid oxidase. Excess amounts can be very toxic. Copper sulfate (blue vitriol) is used as an agricultural poison and as an algicide in water purification. *See* COPPER.

Iron. Iron as a micronutrient is most effective in the ferrous state, $(Fe)^{2+}$. Though it can be absorbed in the ferric state, $(Fe)^{3+}$, iron is reduced to ferrous form in plant tissues. Iron is intimately associated with several enzymes which take part in the oxidative reactions of living cells. It is required as a catalyst in the elaboration of chlorophyll. Iron is not mobile within the tissues, and there is little redistribution of the plant's reserve under conditions of supply deficiency. The availability of iron in the soil is strongly influenced by the pH (or degree of acidity) and is favored by soil conditions that are slightly acidic. The amount of iron found in green plants varies within the range of 10–1500 ppm with deficiency evidenced by chlorosis. Where soil conditions tend to lock iron in less available form, soil or foliant applications of organic chelated iron compounds can relieve the plant deficiency. *See* IRON.

Manganese. Manganese is widely distributed in nature, and only small soil concentrations are required for healthy plant growth. In the manganous form this element is an activator of some enzyme systems, such as carboxylases and dehydrogenases. Plants suffering from manganese deficiency exhibit a chlorosis distinctly different from the symptoms of magnesium or iron deficiency. Soil availability of manganese is favored by slight acidity and much restricted by alkaline conditions. Excess manganese can be very toxic to plants. An excess of this element acts as an oxidizing agent to convert ferrous iron to the ferric state, which is physiologically inactive in plants and thus also contributes to iron deficiency. Manganese is believed to be essential in animals for utilization of vitamin B_1. *See* MANGANESE.

Molybdenum. Molybdenum is the plant nutritional element required in the smallest concentration. The requirement varies among crops, but deficiency in tomato plants is prevented by concentration in the order of 1 part in 100,000,000 in the cultural solution. One of the important functions of molybdenum in plant metabolism is the reduction of nitrates. It is also associated with tannin synthesis and legume nodule formation in nitrogen fixation. *See* MOLYBDENUM.

Zinc. Zinc is an essential growth element for both plants and animals. The amounts of zinc found in whole green plants vary between 3 and 150 ppm. Zinc has a function in relation to the enzyme carbonic anhydrase and is known to be required for the synthesis of indoleacetic acid. Excess soil concentrations can be harmful. Zinc is less available in alkaline soils than those on the acid side, but the effect is less pronounced than it is in the case of manganese or iron. The sulfate is commonly used for fertilizer in zinc-deficient soils. Organic chelates of zinc are sometimes used in foliant applications to ornamental plants, such as gardenias. *See* CHELATION; PLANT MINERAL NUTRITION; ZINC. Alan Longacre

Organic Fertilizers

Organic fertilizers are organic materials of vegetable and animal origin which contain certain macro, secondary, or micro nutrients that can be utilized by plants after application to agricultural soils. Some materials can be applied beneficially to crops in an unprocessed state, or after processing by such technologies as composting, drying, fermentation, anaerobic digestion, co-composting, and chemical stabilization. The primary nutrient sources of vegetable origin are crop residues, green manures, oilseed cakes, seaweeds, and miscellaneous food processing and distillery wastes. Also included in this category is biologically fixed nitrogen from legumes in association with root-nodulating bacteria of the genus *Rhizobium*. Animal sources include animal manures and urine, sewage sludge, septage, latrine wastes, and to a lesser extent materials such as blood meal, bone meal, and fish scraps. Often organic fertilizers are of mixed animal and vegetable origin, such as most farmyard manures, rural and urban composts, and sewage effluents and sludges. *See* NITROGEN FIXATION; SOIL MICROBIOLOGY.

The **table** lists the nitrogen, phosphorus, and potassium content, or fertilizer value, and the carbon:nitrogen ratio of some of the most widely used organic fertilizers. The plant nutrient content for most of these materials is quite low compared with most modern synthetic chemical fertilizers. There is also considerable variation in the nutrient content among these organic materials. Those richest in nitrogen include wastes of animal origin, oilseed processing, and legume green manures. Cereal straws, corn stover (the stalks or nongrain part of the crop), seaweeds, farmyard manures, and composts are among the sources that are generally lowest in nitrogen. Likewise, fertilizers of animal origin are richer in phosphorus than crop residues, seaweeds, composts, and farmyard manures, which have the lowest phosphorus content of the materials listed in the table. Sources highest in potassium include plant residues and animal manures, while sewage sludge and some composts are lowest. The nutrient percentages listed in the table are mean values, which can vary considerably depending on the age and diet of animals, processing and storage conditions, and the soil and climatic conditions under which plants are grown.

Use. Organic fertilizers are often used to supplement chemical fertilizers (or vice versa) or as a soil amendment to improve the physical and chemical properties of soil. In some farming systems, such as organic or biological agriculture, they are the primary nutrient sources applied to field and horticultural crops. When organic fertilizers are used primarily as nutrient sources, their chemical composition is important. However, the availability of their contained nutrients for plants, especially nitrogen, may depend more on the C:N ratio than on the nutrient content itself. Materials with a high C:N ratio (for example, wheat straw and corn stover) generally have a low potential for release, or mineralization, of nitrogen as well as other nutrients to a plant-available form. When applied to soil, such materials may lead to a temporary decrease in plant-available nitrogen because of the concomitant demand for nitrogen by soil microorganisms engaged in residue decomposition. Materials with a lower C:N ratio, such as animal or green manures, would release their nitrogen and other nutrients to crops much more readily with little danger of microbial immobilization.

Types. Thus, organic fertilizers can be divided into general types. First, those which have a high nutrient availability index (NAI) are rapidly decomposed by soil microorganisms, and are essentially unstable organic materials. Examples include fresh animal

Nutrient composition of selected organic fertilizers of vegetable and animal origin*

Fertilizer	C:N	N, %	P, %	K, %
Vegetable				
Wheat	100	0.5	0.1	1.1
Corn stover	55	0.6	0.3	1.3
Green manure (alfalfa)	19	2.5	0.3	2.0
Cottonseed cake	10	6.4	1.3	1.8
Seaweed	17	0.7	0.3	1.2
Animal				
Cattle manure	19	1.9	0.6	1.4
Poultry manure	12	3.8	1.9	1.8
Latrine wastes	8	4.0	1.4	1.3
Bone meal	—	3.4	10.9	0
Blood meal	3.5	11.1	0.7	0
Sewage sludge	10	4.0	1.4	0.3
Fish scraps	4.5	7.5	2.8	0.8
Mixed				
Farmyard manure	15	0.8	0.2	0.7
Composted sewage sludge	19	1.3	0.9	0.2
Composted cattle mature	17	0.4	0.2	0.2

*Most data compiled by the Food and Agriculture Organization (1983).

manures, sewage sludge, and latrine wastes. A second group of materials are characterized by a high organic stability index (OSI), are decomposed relatively slowly in soil, and have a relatively low NAI. An example is composed organic wastes. Materials having a high NAI generally have a low OSI. Farmers in both developed and developing countries utilize both types of materials depending on their need to release nutrients rapidly to a vegetable crop, or to improve the physical properties and productivity of a marginal soil.

Production methods. The value of some organic wastes and residues as organic fertilizers and soil conditioners can often be increased through certain biological or physical processes which enhance consumer demand, storage, handling, and field application. Some of the more common methods include composting, drying, and shredding or grinding.

Compost is produced by the controlled decomposition of organic materials such as manure, sewage sludge, and plant residues by microorganisms under specific conditions of time, aeration, temperature, and moisture. Composting provides organically stable, humuslike products which no longer resemble the input raw materials, and which can be more easily stored, handled, and applied. The process is usually carried out in a compost pit or by heaping the material in stacks aboveground. Composting also reduces bulk and destroys pathogenic microorganisms which may be present in some raw materials.

Drying, usually by exposure to wind and sunlight, is used to remove the large amounts of water present in some materials such as sewage sludges and animal manures. Grinding or shredding can reduce the bulk of coarse materials such as bones and some woody plant residues which are difficult to use in their original form.

Advantages and disadvantages. A distinct advantage of organic fertilizers over concentrated chemical fertilizers is that they release their plant nutrients more slowly over a greater period of time. This slow-release characteristic can greatly minimize the potential for loss through leaching and runoff. Moreover, the organic component is highly beneficial for improving soil physical properties, resulting in increased infiltration and retention of water, soil aeration, and plant growth. Organic fertilizer may also serve as a buffering agent for stabilizing soil pH, and as a complexing agent for reducing phytotoxic effects of aluminum, iron, and manganese under acid soil conditions. Disadvantages of organic fertilizers include their relatively low nutrient analysis, which involves greater bulk and weight than high-analysis chemical fertilizers, and their considerable cost in transportation and labor for application. Continued use of some organic materials as sole sources of fertilizer can lead to an undesirable imbalance of nutrients in the soil. However, where both vegetable and animals sources are applied, there is little likelihood of this occurring. Other potential disadvantages of organic fertilizers are related to public acceptance of some materials (for example, latrine wastes and sewage sludge), health hazards from pathogens and fly breeding, and malodors. However, most of these problems can be overcome.

Quantities available. Organic fertilizers are a major source of plant nutrients for crop production in the United States and in many developing countries. For example, of the approximately 23×10^6 tons (21×10^6 metric tons) of nitrogen returned annually to cropland by United States farmers, 55% originates from organic sources, including biologically fixed nitrogen. This includes 3×10^6 tons (2.7×10^6 metric tons) from crop residues, 1.4×10^6 tons (1.3×10^6 metric tons) from animal manures, and 7.2×10^6 tons (6.5×10^6 metric tons) from biologically fixed nitrogen in legume and green manure crops. Crop residues and animal manures account for about 75% of the organic materials produced annually in the United States, while municipal refuse (garbage or solid waste) accounts for about 18% of the total, and sewage sludge less than 1%. More than 70% of the crop residues and animal manures (up to 35% of the sewage sludge, but only a small percentage of municipal refuse) are utilized as organic fertilizers. It is estimated that the amount of nitrogen, phosphorus, and potassium, respectively, returned to the soil annually in the form of organic fertilizers is equivalent to 85, 55, and 150%, respectively, of the total amount of these nutrients applied as commercial fertilizers. *See* AGRICULTURAL SOIL AND CROP PRACTICES; FERTILIZING; SOIL CHEMISTRY.

R. I. Papendick; J. F. Parr

Bibliography. O. P. Englestad (ed.), *Fertilizer Technology and Use*, 3d ed., 1986; A. Finck, *Fertilizers and Fertilization: Introduction and Practical Guide to Crop Fertilization*, 1982; Food and Agriculture Organization, *Comparative Effects of Mineral Fertilizers and Organic Manures*, Project Field Doc. 14, Regional Project RAS/75/004, FAO/UNDP, New Delhi, 1983; T. P. Hignett (ed.), *Fertilizer Manual*, 1985; J. A. Kent (ed.), *Riegel's Handbook of Industrial Chemistry*, 9th ed., 1992; W. Lockeretz (ed.), *Environmentally Sound Agriculture*, 1983; V. Sauchelli, *Phosphates in Agriculture*, 1965; A. V. Slack and G. R. James, *Ammonia*, pt. 1, Fertilizer Science and Technology Series, vol. 2, 1973; U.S. Department of Agriculture, *Improving Soils with Organic Wastes*, 1978; U.S. Department of Agriculture, *Report and Recommendations on Organic Farming*, 1980; W. H. Waggaman, *o-Phosphoric Acid, Phosphates, and Phosphatic Fertilizers*, 2d ed., ACS Monogr. 34, 1952; R. D. Young and F. P. Acorn, *Trends in U.S. Fertilizer Technology*, Nat. Fert. Dev. Center, TVA, Bull. Y-133, 1978.

Fertilizing

Addition of elements or other materials to the soil to increase or maintain plant yields. Fertilizers may be organic or inorganic. Organic fertilizers are usually manures and waste materials which in addition to providing small amounts of growth elements also serve as conditioners for the soil. Commercial fertilizers are most often inorganic. Fertilizer analysis and systematic application began about 1850 and marked the beginning of scientific crop production.

Grades. Commercial fertilizers are mainly designed to supply one or more of the three major elements nitrogen, N, phosphorus, P, and potassium, K, in suitable chemical form. The fertilizers are graded in the order N-P-K, the numbers indicating the percentage of the total weight of each of the three components. Hence, the numbers 5-10-10 represent a mixture containing 5% nitrogen, N, 10% phosphorus pentoxide, P_2O_5, and 10% potassium oxide, K_2O.

Physical forms. Chemical fertilizers are marketed in both dry and liquid forms. Dry forms include powdered, granulated, and pelleted fertilizers. Liquid fertilizers are obtainable in high-pressure and low- or nonpressure forms. Mixed fertilizer solutions are usually nonpressure liquids.

Application methods. Methods of applying fertilizers vary widely and depend on such factors as kind of crop and stage of growth, application rates, physical and chemical properties of the fertilizer, and soil type. Two basic application methods are used, bulk spreading and precision placement. Time and labor are saved by the practice of bulk spreading, in which the fertilizer is broadcast over the entire area by using large machines which cover many acres in a short time. Precision placement, in which the fertilizer is applied in one or more bands in a definite relationship to the seed or plants, requires more equipment and time, but usually smaller amounts of fertilizer are needed to produce a given yield increase.

When applied at planting, the fertilizer is usually placed in a single band 1–3 in. (2.5–7.5 cm) to the side of and 2–6 in. (5–15 cm) below the seed (**Fig. 1**). Care must be taken to have enough soil between the seed and fertilizer to prevent damage to the seed or developing seedlings. For crops highly susceptible to fertilizer injury, only a small amount of starter fertilizer may be applied at planting, with the remainder being applied as a side-dressing after the plants have emerged. Several side-dressings may be applied during the growing season to those crops with a high requirement for nitrogen or other elements which leach easily. Gaseous or volatile liquid fertilizers are applied 5–6 in. (12.5–15 cm) deep by using a narrow chisel tool and sealing the opening with a press wheel to prevent losses.

Fig. 1. Efficient placement of fertilizer for potatoes. Attachments to make fertilizer placement at the time of planting are available on most field row planters. (*USDA, Agriculture Research Service*)

For some deep-rooted plants, subsoil fertilization to depths of 12–20 in. (30–50 cm) is advantageous. This is usually a separate operation from planting, and uses a modified subsoil plow followed by equipment to bed soil over the plow furrow, thereby eliminating rough soil conditions unfavorable for good seed germination. Top-dressings are usually applied by broadcasting over the soil surface for closely spaced crops such as small grains. Side-dressings for row crops may be placed at depths from surface to 5 in. (12.5 cm), but care must be taken to avoid damage to plant roots from the plow foot.

Equipment. Since solid fertilizers range from dense heavy materials to light powders and liquid fertilizers range from high pressure to zero pressure, a variety of equipment is required for accurate metering and placement. In addition, application rates may be as low as 50 lb/acre (56 kg/hectare) or as high as 6 tons/acre (13.3 metric tons/hectare). Large bulk spreaders usually use drag chains or augers to force the material through a gate or opening whose size is varied to regulate the amount passing through and falling on the spreader. One type of spreader consists of rapidly rotating horizontal disks with vanes mounted on their tops. The fertilizer is spread by centrifugal force over widths varying from 10 to 50 ft (3 to 15 m). Although this method is simple, the accuracy of distribution is affected by particle size, wind velocity, and slope of land on which the application is made. Another spreader consists of long perforated tubes through which the fertilizer is moved by augers.

Often the main supply bin contains simple agitators to prevent heavy or damp materials from bridging. The star-wheel metering unit is commonly used on drills and some row-crop fertilizer hoppers. This dispenser operates horizontally, with fingers carrying increments of fertilizer under a shield which is adjustable in height. Some drills have hoppers with rounded perforated bottoms. The hole size is adjustable by a slide, and the fertilizer is forced through by agitators mounted above the bottom of the hopper. Other mechanisms have been used, but basically all regulate the flow by varying opening size or speed of star wheel or agitator.

Liquid fertilizer of the high-pressure type (for example, anhydrous ammonia) is usually regulated by valves or positive displacement pumps. The size of the orifice may be controlled manually or automatically by pressure-regulating valves. Low-pressure solutions may be metered by gravity flow through orifices, but greater accuracy is obtained by using compressed air or other gases to maintain a constant pressure in the tank. This method eliminates the effect of temperature and volume changes.

Nonpressure solutions may be metered by gravity or by gear, roller, piston, centrifugal, or hose pumps. The accuracy of the gravity type can be improved by the use of a constant head device by which all air is introduced into the tank at the bottom.

All of the pumps except the piston and hose types must be calibrated at a fixed pump speed since their output is not linear. Therefore, the overall accuracy of placement is dependent upon maintaining a constant driving speed, which is very difficult on rolling terrain. Errors of 15% or higher may be introduced by variations in driving speed. The piston and hose pumps can be driven by a wheel rolling over the ground, making their ouput proportional to driving speed. The hose pump consists of any number of hoses which are compressed by rollers in a sequential action which squeezes the liquid along the hoses. Since the output is determined only by the speed at which the pump is rotated, no manifolds, valves, or orifices are needed. Nonvolatile fertilizer solutions are often pumped into the supply lines of irrigation systems to allow simultaneous fertilization and irrigation. With the exception of bulk spreaders and other broadcasters, most fertilizer application devices are built as attachments which can be mounted in conjunction with planters, cultivators, and herbicide applicators. Often the tanks, pumps, and controls used for liquid fertilizers are also used for applying other chemicals such as insecticides.

J. G. Futral

Effect on growth and yield. For plants to develop, all the nitrogen and mineral elements essential for growth must be supplied by way of the root system or, to a limited extent, through the leaves. The total supplies of mineral elements in certain soils are often adequate for many years of crop production, but the rate at which those elements become available for plant use may be too slow. Each nutrient must be in adequate supply in intensity and capacity, and in a reasonably favorable balance with all the others. Hence potassium, phosphorus, and any of several micronutrients may be deficient in rate of supply because of fixation or adsorption by the soil; lack of nitrogen may become a limiting factor through adverse microbial conditions or excessive loss by leaching. Moreoever, the rate of supply of nutrients must be favorable throughout the various ontogenetic stages of growth.

Desirable characteristics. Fruit or storage organs, rather than merely vegetative yield, are often desired. Further, it is frequently the total content of some substance within the harvested organ that is important rather than the weight or size of the organ alone. Thus the quality of tomatoes or grain or the percentage of sugar in cane or beets is economically important. The mineral supply at later stages of growth may determine the economic value of the crop. For example, high nitrogen acquisition at later development stages may lead to increased protein but decreased carbohydrate content. Thus mineral nutrition must be in favorable accord with development of desirable characteristics of vegetative growth and the harvested product.

Nutritional requirements. Plant growth and the quality of the harvested product are dependent on proper nutrition. Mineral materials are usually derived from the soil and, while total supply is important, the rate of effective availability is paramount. Plant and soil form an integrated system, in which the root is the nutrient-absorbing organ. While the type of root development is a species characteristic, its distribution in space may depend on environmental conditions in the medium. Thus it may vary in lateral and vertical disposition and in production of root hairs. Modifications also depend on physical and chemical conditions in the soil. Cultural management, such as cultivation, water, fertilization, spacing (competition), and cropping regime, may be determinative. If rooting into areas of nutrient adequacy is extensive (other factors being satisfactory), additional fertilizer may not be essential. Soil depletion, however, requires resupply to meet plant demand; deficiencies must be corrected for optimum growth. Growth is dependent on nutrient supply. Conversely, since plant need varies with rate of growth and the stage of ontogenetic development, the time, placement, and rate of fertilization are important. Some typical fertilizer placements are illustrated in **Fig. 2**. Plant growth, nitrogen distribution and total sorption in corn are shown in **Fig. 3** and in the **table**.

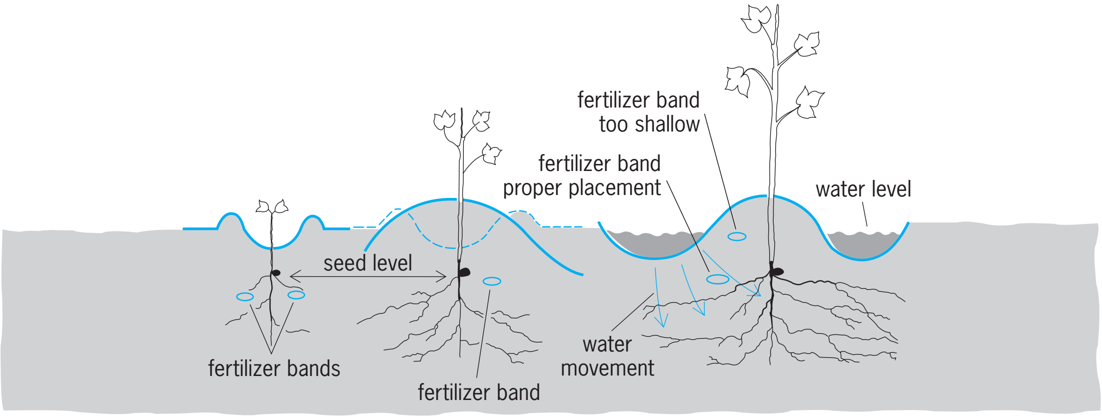


Fig. 2. Relationship of the placement levels of plant food materials to the movement of water. (*After Western Fertilizer Handbook, California Fertilizer Ass., 1965*)

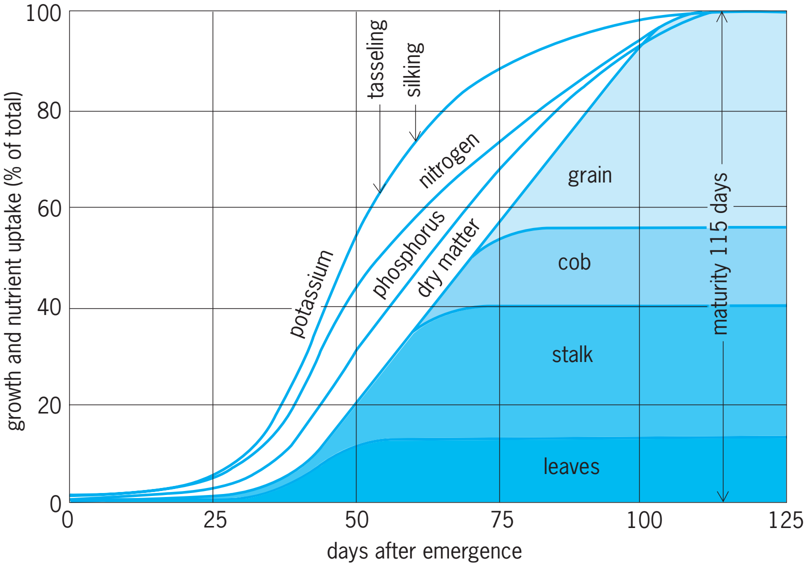


Fig. 3. Graph showing the growth and the uptake of nutrients in relation to the accumulation of dry matter by corn plants. (*After C. E. Millar, L. M. Turk, and H. D. Foth, Fundamentals of Soil Science, Wiley, 1965*)

Nonlegumes may require greater nitrogen fertilization than legumes. In this respect, the microbiological population and other soil factors, such as temperature, aeration, pH, and organic matter, are important. Bacterial species and their numbers are an important facet of fertility relations in general, especially insofar as nitrogen fixation and the presence of organic matter are concerned. Pathological disease incidence varies with fertilization. Fertility relations affect evapotranspiration by modifying the ratio of canopy cover to soil surface area, thus determining efficiency of water use. Rate of replenishment of nutrients from the soil or other medium may limit growth, as root sorption may create depletion zones in the rhizosphere of roots. *See* NITROGEN FIXATION; SOIL MICROBIOLOGY.

Fertilizers. For improving the mineral nutrition of plants, single materials or favorable combinations of chemicals are compounded and used. Chemical compounds may include all of the mineral elements that are essential to growth. Micronutrients, occurring as impurities in fertilizer-grade chemicals, in some cases may be inadequate for maximum plant development and extra supplies may have to be added. Some latitude is permissible in the balance of elements supplied, but both marked deficiencies or marked excesses of any one element must be avoided.

Balanced fertilizers. Commercial fertilizers, manures, lime, or other materials may be added to provide essential mineral elements at required levels of availability or as amendments to make environmental conditions more favorable, improve structure, or enhance microbial activity within the soil so that such factors do not become limiting to growth and development. Sometimes reference is made to the use of balanced fertilizers. The balance that is important is not in the fertilizer but in the soil after the fertilizer has been added and has reacted with the soil. The same fertilizer may produce different effects in different soils. Addition to the soil does not necessarily ensure availability; certain elements, such as potassium, phosphorus, and at least some of the micronutrients, are strongly fixed in some soils. This is especially true where the colloidal clay content of the soil is relatively high. In any case, what is important is the ability of the plant to acquire essential nutrients in proper quantities favorable for the species at any particular time during growth and development, assuming other factors to be optimal.

Soil type and composition. When considering application of fertilizer to soil, general information should be sought regarding the types of soil in the area and their physical and chemical compositions. Crops favorable to the representative soil and climatic conditions should be considered. Local agricultural information on these matters is usually available. "Feeding" zones vary with species and soil management. Shallow-rooted species absorb from and deplete zones different from those of deep-rooted crops. As roots grow, absorb, and die and after root decomposition, minerals may be redistributed thereby in the soil and thus provide a more available source of nutrients for succeeding plantings. Fallowing allows replenishment from parent and residual organic materials. Total salt and pH tolerance of species, as well as the total specific salt content and pH of the soil, must be considered. Protracted applications of single or even some mixed fertilizers may correct unfavorable conditions or cause adverse changes in total salt level, exchangeable ion ratios, or pH of the soil. The

Approximate plant growth, sorption of primary nutrients, and water use for a 100-bu corn crop per acre or hectare

			Nutrient sorption*			
Plant part	Basis	Dry weight	N	P	K	Water use†
Shoots	Acre	13,000 lb	155 lb	26.3 lb	104 lb	6.5×10^4 lb
	Hectare	14,600 kg	173.7 kg	29.5 kg	116.5 kg	7.3×10^4 kg
Roots	Acre	5,000 lb	50 lb	8.7 lb	33 lb	2.5×10^4 lb
	Hectare	5,600 kg	56.0 kg	9.7 kg	37.0 kg	2.8×10^4 kg
Shoots and roots	Acre	18,000 lb	205 lb	35 lb	137 lb	5.5×10^6 lb
	Hectare	20,200 kg	229.8 kg	39.2 kg	153.4 kg	6.2×10^6 kg

*Apart from certain possible luxury accretion, the total sorption of nutrients represents that necessary during growth. That part in roots is restored to the soil. Therefore the net removal of nutrients is that contained in the shoots, unless stover is returned to the soil.

†Total water use during growth (5.5×10^6 lb or 6.2×10^6 kg) includes that for growth and that expended in evapotranspiration. The water in shoots is in large part lost to the air on crop maturity; that in roots is largely returned to the soil.

nature of the materials and of all interrelated facets of cropping and soil dynamics determine the result. For example, simple leaching may correct total salinity; sulfur, with microorganism activities, may correct alkalinity; lime may correct acidity; and gypsum may correct an unfavorable ion-exchange balance and the physical structure of soils.

Types of fertilizers. Fertilizer materials may be organic (urea) or inorganic (ammonium phosphate, potassium nitrate); natural (Chilean sodium nitrate, calcium sulfate) or synthetic (ammonium sulfate); simple, mixed (two or more), or complete (containing N, P, and K). They may be applied to soil in gaseous (ammonia), liquid (ammonia, or solutions), or solid (rock phosphate) states. Filler or make-weight material may be added to solid concentrates. Numerous factors must be considered in deciding the most productive mode, place, amount, balance, and time of application to fulfill known or anticipated plant needs. Knowledge of specific crop requirements at various stages of development (seedling, vegetative, reproductive), soil and irrigation water analysis, periodic plant tissue tests, and previous growth and harvest results will assist in suggesting management practice.

Maintaining adequate soil conditions. Natural and artificial materials can replenish nutrients as well as improve structure and water-holding capacity of many soils. Temperature, pH, aeration, and activity of microorganisms determine the rate of decomposition and supply or loss of nutrients. Natural manures are good sources of N, P, and K; stubble may be low in N, P, and Ca, but may be supplemented with ammonium sulfate, superphosphate, and limestone. Cover crops, especially legumes, are a means of effectively supplying natural organic matter to soils and favorably altering nutrient availability. Similarly, crop rotation can usefully provide protracted land use while meeting economic supply and demand. It is often necessary on a long-term basis to obviate or correct development of certain soil pathogens. Nurse crops under some conditions (for example, in orchards) can be beneficial. Manures, though beneficial on many soils, are not necessarily indispensable. Fully satisfactory plants or their products can be attained with soilless culture. Although soilless culture cannot supplant usual agricultural methods generally, it can supplement them where mechanics and economics justify such methods.

Choice of fertilizer. The best fertilizers to use depend on the anticipated plant need and interactions with the soil, prior to or during cropping, which may determine availability for plant growth. Organics and many inorganics are useful where suitable. Need must be known, and the appropriate combinations of elements must be applied to favor soil conditions and plant growth. A single inorganic salt or combinations of such salts may accomplish this. The ratios of essential elements in a fertilizer should be considered on their atomic equivalent basis per unit weight, not as the oxides or otherwise, which is the way the plant utilizes them. Further, size of particle and nature of the chemicals should be considered; granules or larger particles and relatively insoluble compounds may dissolve slowly, which may be desirable or disadvantageous. Some applied ions may be fixed on soil colloids (such as K or P) or precipitated (such as Ca, Fe, or P) through chemical double decomposition. Some solubles (such as NO_3) may be leached by excessive amounts of water to subsoil or drainage systems and thus become ineffective or lost.

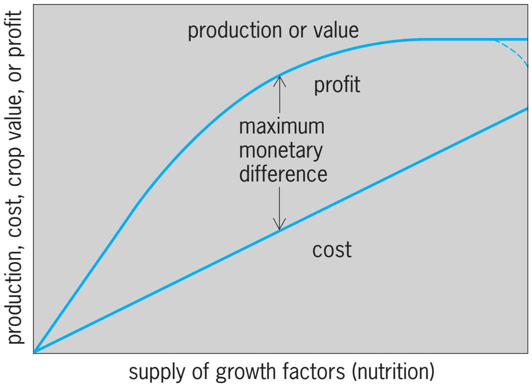


Fig. 4. Representation of the economics of crop production per unit area. Schematic shapes of curves (increases) may vary with conditions.

So-called sustained-release fertilizers are marketed; they are coated particles that release the elements slowly and protractedly. To lessen fixation of micronutrient cations by soil colloids, these metals are effectively supplied in chelated form, which involves weak chemical binding. In this form they may be sprayed upon foliage and thus effectively fertilize the plant. Natural materials may supply micronutrients as an appreciable impurity, though often inadequately.

Soil conditioning. Soil amendments or conditioners are often applied to improve structure for aeration, water penetration, and other soil conditions. Natural organic materials may be helpful, if not supplied excessively. Gypsum can favorably increase granulation and flocculation of soil colloids, being especially useful in nonsaline alkali soils. Some synthetics have been proposed to improve soil structure but thus far are not economically practical. Some have been suggested to decrease excessive evapotranspiration but, here too, further experimental trial is necessary for practical and economic feasibility.

To properly fertilize a crop, the agriculturist should know fairly well the nature and properties of soils, fertilizers, and plants, their many interactions, and the economics of fertilization (**Fig. 4**). *See* AGRICULTURAL MACHINERY; AGRICULTURAL SOIL AND CROP PRACTICES; FERTILIZER; PLANT GROWTH.

Theodore C. Broyer

Bibliography. J. Archer, *Crop Nutrition and Fertilizer Use*, 2d ed., 1988; California Fertilizer Association, *Western Fertilizer Handbook*, 8th ed., 1994; G. W. Cooke, *Fertilizing for Maximum Yield*, 3d ed., 1988; O. P. Engelstad (ed.), *Fertilizer Technology and Use*, 3d ed., 1985; S. L. Tisdale and W. L. Nelson, *Soil Fertility and Fertilizers*, 6th ed., 1998; K. K. Vyas and S. W. Saxena, *Biofertilizers*, 1990.

Fescue

A group of approximately 100 species of grass; more than 30 are represented in the United States. Tall fescue (*Festuca arundinacea*), a perennial cool-season plant introduced from Europe, occupies about 35×10^6 acres (15×10^6 hectares), primarily in the humid south-central region of the United States. It is popular because of its ease of establishment, vigor, wide range of adaptation, long grazing season, tolerance to abuse, sufferance of drought and poor soils, pest resistance, good seed production, and esthetic value when used for turf, ground cover, and conservation purposes. Its dominance of pastures and roadsides indicates that it fills an important ecological niche. It is used primarily as pasture and hay for beef cattle, with lesser use for dairy cows or replacement heifers, sheep, and horses. The leafy and vigorous plants can grow to 3–4 ft (0.9–1.2 m) if undisturbed; under grazing or clipping, they can form a dense sod when sufficient water and fertility are available.

Tall fescue toxicosis. Cattle and sheep that graze tall fescue can suffer from a condition known as tall fescue toxicosis. During warm weather, signs of this toxicosis are poor animal gains [less than 450 g/day (16 oz/day) for steers], intolerance to heat, excessive salivation, rough haircoat, slightly elevated body temperature, nervousness, lowered milk production, and abnormally low conception rate. In cool weather, the toxicosis can be expressed as fescue foot, a gangrenous condition of cattle feet. Bovine fat necrosis, characterized by fatty deposits in cattle abdomens, is associated with applications of several tons of chicken manure per hectare to tall fescue pastures. When pregnant mares graze tall fescue, particularly during the last trimester of gestation, their placentas become thickened, which sometimes causes delivery of dead foals, little or no milk production even when the foal survives birth, and sometimes death of the mare.

An endophytic symbiotic fungus, *Acremonium coenophialum*, growing between the cells of tall fescue tissues, is responsible for the toxicosis. Over 80% of all tall fescue plants contain the endophyte; that is, they are E^+.

Cultivars. There are over two dozen cultivars of tall fescue, the most widespread being Kentucky 31 in variants that either contain the fungus endophyte or are endophyte free. The use of endophyte-free cultivars to improve animal performance is important. The turf industry uses cultivars that contain the endophyte for their greater tolerance to environmental stress; special storage is used to guarantee viability of both endophyte and seed.

Other important fescues include meadow fescue (*F. elatior*), used to a limited extent in pasture; red fescue (*F. rubra*) and Chewings fescue (*F. rubra* var. *commutata*), a fine-leaved, shade-tolerant species with several cultivars used for turf; and Idaho fescue (*F. idahoensis*) and sheep fescue (*F. ovina*), fine- or bristle-leaved species adapted to drier regions and poorer soils. *See* CYPERALES; GRASS CROPS.

Henry A. Fribourg

Bibliography. C. W. Bacon and M. R. Siegel, Endophyte parasitism of tall fescue, *J. Prod. Agr.*, 1:45–55, 1988; D. M. Ball, J. F. Pedersen, and G. D. Lacefield, The tall-fescue endophyte, *Amer. Sci.*, 81(4):370–379, 1993; H. A. Fribourg, C. S. Hoveland, and K. D. Gwinn, Tall fescue and the fungal endophyte: A review of current knowledge, *Tenn. Farm Home Sci.*, 160:30–37, 1991.

Fetal alcohol spectrum disorder

A range of effects, including physical, mental, behavioral, and/or learning disabilities, that can occur in the offspring of women who consume alcoholic beverages during pregnancy. The term fetal alcohol spectrum disorder encompasses several conditions. Fetal alcohol syndrome is the most severe disorder, and is characterized by abnormal facial features and growth and central nervous system problems.

The frequency of adverse pregnancy outcome of chronic alcoholic women is around 40%. There is an eightfold increase in perinatal mortality, and a smaller percentage has the dysmorphic features noted above.

Characteristics. Children who are born to women who consume alcohol during pregnancy may exhibit pre- and postnatal growth deficiency. At age 1, they often continue to be small and fail to thrive. Similarly, some children score below average on standard intelligence and motor function tests. Physical abnormalities include typical facial changes, such as small eye openings and a lack of distinction (or a groove) in the area from below the nose to the upper lip. The presence of these characteristic abnormal facial features is sufficient to alert physicians to possible unrecognized alcoholism in the mother and to enable the diagnosis of fetal alcohol syndrome during the neonatal period, although the typical facial changes are more fully recognizable after age 3.

The severity of the dysmorphic features appears to correlate with the extent of cognitive dysfunction. With age, more manifestations of complex and pervasive neurobehavioral deficits become apparent. These include cognitive dysfunction, attention deficit and hyperactivity, learning disability, poor adaptive behavior, and conduct disorder. The spectrum of deficits can vary widely. Fetal alcohol spectrum disorder encompasses all these manifestations. *See* ATTENTION DEFICIT HYPERACTIVITY DISORDER.

Level of alcohol consumption. Although full-blown fetal alcohol syndrome is seen in some children born to severely alcoholic mothers, lesser degrees of change may be found in some offspring of women who consume smaller amounts of alcoholic beverages. It has yet to be determined whether there is a threshold dose of alcohol consumption during pregnancy below which no fetal injury results. However, fetal alcohol spectrum disorder has been described only among offspring of problem-drinking mothers.

It is not known whether there is a safe level of alcohol consumption during pregnancy. Even at a level of two drinks per day, infants average about 2–5.5 oz

(60–160 g) below the expected birth weight for gestational age. Whether small amounts of alcoholic beverages produce injury or death of the embryo during these early stages of development remains to be established. Women should be advised to refrain from any alcoholic beverages during gestation.

Pathogenesis. The mechanisms by which ethanol produces the changes associated with fetal alcohol spectrum disorder in the developing fetus are not known. One potential mechanism implicates the major ethanol metabolite, acetaldehyde, which may directly affect the proliferation of fetal cells. Alcoholic women tend to have higher levels of circulating acetaldehyde, probably due to their development of alcohol-related liver disease. However, there is other evidence to suggest that ethanol itself can produce the changes noted. New evidence points to oxidative damage. Antioxidants can ameliorate experimentally induced brain injuries caused by ethanol administration. *See* ANTIOXIDANT.

Whether the concurrent use of alcohol and tobacco or the common nutritional deficiencies associated with alcoholism have a potentiating role remains to be determined. The impairment of growth and development and cognitive functions are possible following the maternal use of a number of central nervous system–active agents. Thus, the syndrome described may be augmented by a variety of agents or may be the result of a common mechanism induced by a variety of agents. *See* ALCOHOLISM; BEHAVIOR TOXICOLOGY.

Prevention. Medical advice regarding the consumption of alcoholic beverages during pregnancy involves several principles: heavy drinking is clearly deleterious to the fetus; abstinence from alcoholic beverages immediately preceding and during pregnancy is the only assured safe course of action presently known; and therapeutic attention should also be directed to the problems leading to the alcoholism so that the psychological environment of the offspring may be improved.

Testing for the presence of fatty acid ethyl esters (ethanol metabolites) in meconium (first neonatal stool) is becoming a clinical tool to identify babies exposed in utero to extensive maternal drinking and thus identify infants at risk of developing fetal alcohol spectrum disorder.

Gideon Koren; Irena Nulman; N. Karle Mottet

Bibliography. E. L. Abel, *Fetal Alcohol Syndrome*, 1990; D. Chan et al., Fetal exposure to alcohol as evidenced by fatty acid ethyl esters in meconium in the absence of maternal drinking history in pregnancy, *Ther. Drug Monit.*, 26:474–481, 2004; S. Clarren et al., Brain malformations related to prenatal exposure to alcohol, *Pediatrics*, 92:64–67, 1978; K. Jones and D. W. Smith, Recognition of the fetal alcohol syndrome in early infancy, *Lancet*, 2:999–1001, 1973; G. Koren and I. Nulman, *Motherisk Handbook for the Diagnosis of Fetal Alcohol Syndrome*, Motherisk Program, Toronto, 2002; G. Koren et al., Fetal alcohol spectrum disorder, *CMAJ*, 169:1181–1118, 2003; A. P. Streissguth et al., Teratogenic effects of alcohol in humans and laboratory animals, *Science*, 209:353–361, 1980.

Fetal membrane

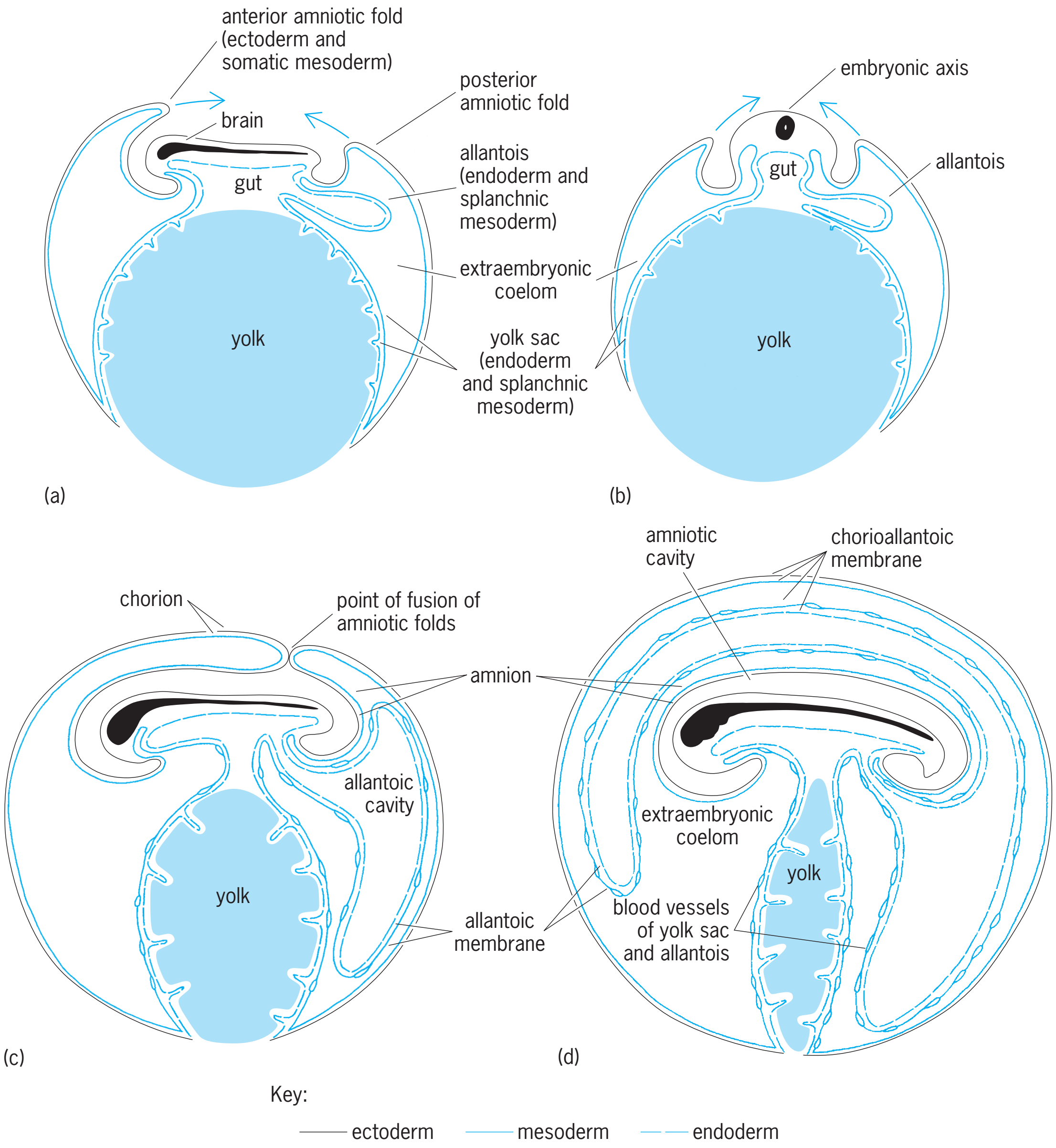
Any of four membranous structures surrounding the embryo during its developmental period. Because these membranes are external to the embryo proper, they are called extraembryonic membranes. Throughout embryonic development, every embryo must contend with a few basic needs. All developing animals need some source of nutrition, a way to remove wastes, some form of respiration, and protection from the outside environment. Fetal membranes collectively serve these functions.

Amniote egg. Invertebrates and lower vertebrates usually undergo rapid development in an aqueous environment until the free-swimming, feeding larval stage. Their eggs are characterized by membranes that are produced by the maternal follicle cells or by secretions from specialized cells of the maternal reproductive tract. As animals made the transition to life completely on land, they adapted a mechanism to enclose the developing embryo in a self-maintained aqueous state. This led to the formation of the amniote egg. Reptiles, birds, and mammals (collectively known as amniotes) undergo a longer developmental period than fish and amphibian embryos and thus require greater protection from the external environment. Whether they are encased in a shell or the uterus, amniotic embryos form four extraembryonic membranes—yolk sac, allantois, chorion, and amnion—to assist their development in the terrestrial environment. These fetal membranes are formed by the tissues of the developing embryo itself and, therefore, differ from membranes or shells that are of maternal origin. *See* AMNIOTA.

Development and structure. In early development, there is no initial delineation between embryonic and extraembryonic structures. The fetal membranes are made up of the same sheets of cells as the embryo: the ectoderm, the endoderm, and the lateral plate mesoderm. As development progresses and the body folds form by differential proliferation rates, the body becomes separate from the yolk and thus embryonic and extraembryonic structures become defined. The extraembryonic ectoderm and mesoderm combined (called the somatopleure) form the amnion and the chorion. The extraembryonic endoderm and mesoderm combined (called the splanchnopleure) form the yolk sac and allantois. The space between these two mesodermal layers is continuous with the body cavity (coelom) of the embryo and is called the extraembryonic coelom. *See* COELOM; GERM LAYERS.

Yolk sac. The yolk sac is the first of the four extraembryonic membranes to form. It is of endodermal origin and is connected to the midgut by the yolk duct. As the embryo develops, the nutrients within the yolk sac are used up and it decreases in size, ultimately becoming incorporated into the gut of the embryo.

Chorion and amnion. The amnion and chorion initially form over the dorsal region of the embryo, but then extend ventrally and fuse to become the external layers. To do this, the ectoderm plus the somatic mesoderm form a double-layered fold (the anterior amniotic fold) anterior to the embryonic axis



Relationships of fetal membranes to each other and to the developing embryonic body. (*a*) Midsagittal section during amniotic folding. (*b*) Cross section showing lateral amniotic folds. (*c*) Midsagittal section just before fusion of amniotic folds. (*d*) Midsagittal section showing fully developed fetal membranes.

(**illus.** *a*). A similar fold (the posterior amniotic fold) appears posterior to the axis a short time later. As these folds grow toward each other over the dorsal side of the embryo, lateral folds also appear and approach each other (illus. *b*). The result is that the exposed surface of the embryo gradually disappears beneath the advancing margins of the folds (illus. *c*), which soon meet and fuse over the posterodorsal region of the embryonic axis. Following fusion of the two sides, the upper double-layered membrane (now the chorion) separates from the lower (now the amnion), and the extraembryonic coelom becomes continuous over the dorsal surface of the embryo. The chorion becomes the outermost, and the amnion the innermost, membrane surrounding the developing embryo (illus. *d*).

Allantois. In a similar way, the allantois begins as a blind sac growing out from the ventral wall of the hindgut (and is, therefore, lined with endoderm), pushing ahead a covering of splanchnic mesoderm (illus. *a*, *b*, and *c*). As the allantois increases in size, it expands into the extraembryonic coelom, and its splanchnic mesoderm becomes closely associated with, if not fused with, the somatic mesoderm of the chorion. The two membranes together are called the chorioallantoic membrane (illus. *d*).

Function. Each of the four fetal membranes serves a different primary purpose: the yolk sac contains the source of nutrients, the allantois is responsible for waste management, the chorion allows for gas exchange, and the amnion produces and maintains the amnionic fluid, which protects the embryo from drying out.

Yolk sac. The primary function of the yolk sac is for the external storage and supply of nutrients. Instead of the nutrients being acquired directly through the yolk duct, nutrients are obtained through the blood vessels found in the mesoderm of the splanchnopleure around the yolk sac. In placental mammals, the yolk sac is not necessary for nutrient storage, as the mother supplies what the embryo needs through the placenta; however, a vestigial yolk sac remains. This is possibly due to the secondary functions it performs: the yolk sac endoderm is the source of primordial germ cells, as well as the source of original red and white blood cells, which originate from mesodermal cells lining the yolk sac endoderm. This endoderm also contributes in development as an inducer of specific organs and organ systems.

Allantois. The primary function of the endodermally lined allantois is the storage of urinary waste. This is most important in reptiles, birds, and nonplacental mammals undergoing external development, as it limits the embryo's exposure to toxins and osmotic stress. In birds and reptiles, this waste is primarily in the form of uric acid, which is relatively insoluble in water and must be stored away from the embryo in the allantois. The close association of this membrane (in conjuction with the chorion) with the shell also allows for some level of gas exchange in these organisms. Among mammals, there is great species variation, and in each species the level of allantois involvement is inversely related to the amount of vascular exchange occurring in the placenta. Thus, the allantois also has some function in gas exchange.

Chorion. The chorion, the outermost of the four membranes, is the site of exchange with the environment. In bird and reptile eggs the chorion's principal function is gas exchange, whereas in mammals the chorion develops into the placenta, and its functions include respiration, nutrition, excretion, filtration, and endocrine and immune responses. Depending on the species of placental mammal, the placenta may be easily separated from the uterus (as in pigs) or may be deeply integrated with the maternal components of this tissue (as in humans). *See* PLACENTATION.

Amnion. The amnion is responsible for the production and maintenance of the amniotic fluids. This aqueous environment is essential to prevent the embryo from drying out, to allow for proper folding of the developing embryo and membranes, and to allow for exchange of gases and waste. Smooth muscle fibers in the amnion spontaneously contract and gently rock the embryo before it develops the capacity for movement.

Fate. Although these four fetal membranes are essential to development, very little of them remains once the animal is born or hatched. The yolk sac is incorporated into the midgut, and the proximal portion of the allantois becomes the urinary bladder. The amnion and chorion separate entirely from the animal, and the placenta (including part of the maternal uterine mucosa, in some cases) is delivered separately as afterbirth. *See* ALLANTOIS; AMNION; CHORION; YOLK SAC. Cheryl Schaeberle

Bibliography. B. I. Balinsky, *An Introduction to Embryology*, 5th ed., 1981; F. C. Battaglia and G. Meschia, *An Introduction to Fetal Physiology*, 1986; B. M. Carlson, Extraembryonic membranes and placenta, in *Patten's Foundations in Embryology*, 5th ed., Chap. 7, pp. 250–289, McGraw-Hill, New York, 1988; J. C. Cross, Formation of the placenta and extraembryonic membranes, *Ann. N.Y. Acad. Sci.*, 857:23–31, 1998; S. F. Gilbert, *Developmental Biology*, 7th ed., Sinauer Associates, Sunderland, MA, 2003; N. K. Wessels, *Tissue Interactions and Development*, 1977.

Fever

An elevation in the central body temperature of warm-blooded animals caused by abnormal functioning of the thermoregulatory mechanisms. Fever accompanies a wide variety of disease states, both infectious and noninfectious, and in the great majority of instances is due to an abnormality in the regulation of body temperature by the central nervous system. *See* THERMOREGULATION.

Endogenous pyrogens. Experimental studies on the cause of fever suggest that leukocytes (circulating granulocytes and monocytes) of the host, as well as the fixed macrophages of the reticuloendothelial system in the liver and spleen, can be activated by various stimuli to produce a fever-inducing substance—endogenous pyrogen. Activators of microbial origin include the lipopolysaccharide endotoxins of gram-negative bacteria, cell walls and soluble antigens of other bacteria, viruses, and fungi. Endogenous pyrogen has been characterized as a protein of relatively low molecular weight (13,000) with an essential lipid moiety. Pyrogens from different tissues and from rabbits and humans appear to have similar characteristics. Circulating endogenous pyrogen, liberated by activated cells, is believed to cross the blood-brain barrier and act upon specialized neurons, presumably located in the anterior hypothalamus, that initiate heat conservation and increase heat production.

Mechanisms. The mechanisms responsible for elevating body temperature include: reduction in heat loss by constriction of peripheral vessels whose tone is under control of the sympathetic nervous system; inhibition of panting and sweating, the latter by way of the cholinergic nerves; and increased heat production by means of shivering in voluntary muscles innervated by somatic motor nerves.

Pharmacologic pyrogens. Various naturally occurring pharmacologic agents, including serotonin (5-hydroxytryptamine) and catecholamines, are normally present in high concentration in areas of the brain that control body temperature and can cause fever when injected under certain experimental circumstances. The role of these agents in most experimentally produced or clinical fevers is obscure. Certain naturally occurring steroids are also pyrogenic and are known to activate human leukocytes in the test tube.

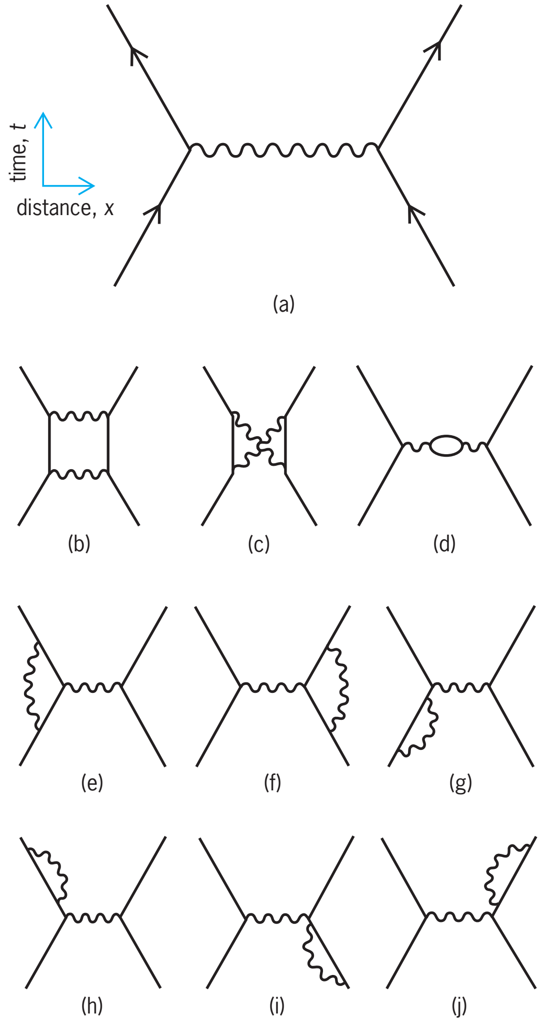
Effects. There is no clear evidence that elevated body temperature evoked by most infections is directly injurious to microbial invaders. As fever regularly accompanies inflammation, however, increased body temperature may well accelerate certain biochemical reactions of use to the host in combating infection. *See* HOMEOSTASIS; INFLAMMATION. Elisha Atkins

Bibliography. H. Laudien, *Temperature Regulations in Animals and Man*, 1986; J. B. Walter, *Pathology of Human Disease*, 1989; B. W. Zweifach et al. (eds.), *The Inflammatory Process*, 2d ed., 1974.

Feynman diagram

A pictorial representation of elementary particles and their interactions. The diagrams show the paths of particles in space and time as lines, and the interactions between the particles as points where the lines meet.

These diagrams were introduced by R. P. Feynman in 1949 in the context of quantum electrodynamics, the quantum field theory of electromagnetic interactions of charged particles. The **illustration** shows Feynman diagrams for electron-electron scattering. In each diagram, the straight lines represent space-time trajectories of noninteracting electrons, and the wavy lines represent photons, particles that transmit the electromagnetic interaction. External lines at the bottom of each diagram represent incoming particles (before the interactions), and lines at the top, outgoing particles (after the interactions). Interactions between photons and electrons occur at the vertices where photon lines meet electron lines. *See* ELECTRON; PHOTON.



Feynman diagrams for electron-electron (Møller) scattering: (*a*) second-order diagram (two-vertices); (*b–j*) fourth-order diagrams.

Each Feynman diagram corresponds to the probability amplitude for the process depicted in the diagram. This amplitude can be calculated from the structure of the diagram according to specific rules. It is a product of wave functions for external particles, a propagation function for each internal electron or photon line, and a factor for each vertex that is proportional to the strength of the interaction, summed over all possible space-time locations of the vertices. The propagation function, corresponding to a line joining two vertices, is the amplitude for the probability that a particle starting at one vertex will arrive at the other. *See* PROPAGATOR (FIELD THEORY).

The set of all distinct Feynman diagrams with the same incoming and outgoing lines corresponds to the perturbation expansion of a matrix element of the scattering matrix in field theory. This correspondence can be used to formulate the rules for writing the amplitude associated with a particular diagram. The perturbation expansion and the associated Feynman diagrams are useful to the extent that the strength of the interaction is small, so that the lowest-order terms, or diagrams with the fewest vertices, give the main contribution to the matrix element. *See* PERTURBATION (QUANTUM MECHANICS); SCATTERING MATRIX.

Since their introduction in quantum electrodynamics, Feynman diagrams have been widely applied in other field theories. They are employed in studies of electroweak interactions, certain situations in quantum chromodynamics, acoustooptics, and in many-body theory in atomic, nuclear, plasma, and condensed matter physics. *See* ACOUSTOOPTICS; ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; QUANTUM CHROMODYNAMICS; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; WEAK NUCLEAR INTERACTIONS. Peter Mohr

Bibliography. J. Bjorken and S. Drell, *Relativistic Quantum Fields*, 1965; J. Bjorken and S. Drell, *Relativistic Quantum Mechanics*, 1964; C. Itzykson and J. Zuber, *Quantum Field Theory*, 1980; M. Kaku, *Quantum Field Theory: A Modern Introduction*, 1993; F. Mandl and G. Shaw, *Quantum Field Theory*, paper 1994; M. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, 1995; P. Ramond, *Field Theory: A Modern Primer*, 2d rev. ed., 1990.

Feynman integral

A technique, also called the sum over histories, which is basic to understanding and analyzing the dynamics of quantum systems. It is named after fundamental work of R. Feynman. The crucial formula gives the quantum probability density for transition from a point q_0 to a point q_1 in time t as expression (1), where S(path) is the classical mechanical action

$$\int \exp\left[iS(\text{path})/\hbar\right]\, d\left(\text{path}\right) \qquad (1)$$

of a trial path, and $\hbar$ is the rationalized Planck's constant. The integral is a formal one over the infinite-dimensional space of all paths which go from q_0 to q_1 in time t. Feynman defines it by a limiting procedure using approximation by piecewise linear paths. *See* ACTION.

Feynman integral ideas are especially important in quantum field theory, where they not only are a useful device in analyzing perturbation series but are also one of the few nonperturbative tools available.

An especially attractive element of the Feynman integral formulation of quantum dynamics is the classical limit, $\hbar \to 0$. Formal application of the method of stationary phase to expression (1) says that the significant paths for small $\hbar$ will be the paths of stationary action. One thereby recovers classical mechanics in the hamiltonian stationary action formulation. *See* LEAST-ACTION PRINCIPLE.

Feynman-Kac formula. Mark Kac realized that if one replaces the time by a formal purely imaginary time (that is, replaces the Schrödinger equation by a heat equation), then one can make precise mathematical sense out of Feynman's integral in terms of the stochastic process introduced by Norbert Wiener in his theory of brownian motion. The resulting formula for the semigroup e^{-tH}, with $H = -{}^1\!/_2\Delta + V$ the quantum hamiltonian (Δ is the laplacian operator and V is the potential energy), is given by Eq. (2).

$$(e^{-tH}f)(x) = E_x\left[\exp\left(-\int_0^t V[b(s)]\,ds\right) f[b(t)]\right] \qquad (2)$$

Here E_x is the expectation with respect to Wiener's process, $b(s)$, starting at the point x, and f is an arbitrary bounded measurable function to which e^{-tH} is applied. Equation (2) is the Feynman-Kac formula. It makes a variety of probabilistic ideas applicable to a rigorous mathematical analysis of quantum theory. *See* BROWNIAN MOVEMENT; STOCHASTIC PROCESS.

Euclidean quantum field theories. The passage to imaginary time initiated by Kac is especially useful in relativistic quantum field theory. Since the Lorentz covariance of relativity is replaced by euclidean group covariance under this replacement, the new theory is often called euclidean field theory. The corresponding Feynman integral is often called a euclidean path integral and has become the standard formulation of Feynman integrals in most modern discussions of quantum field theory. The analogs of stationary action solutions are called instantons. *See* INSTANTON.

Statistical mechanics analogy. The appearance of the exponential in Eq. (2) suggests an analogy between quantum dynamics and classical statistical mechanics. The analog of the total space-time dimension in a euclidean quantum field theory is then the dimension of space in statistical mechanics. Just as the phenomenon of phase transitions occurs in statistical mechanics, the analogous phenomena of spontaneously broken symmetries and dynamical instability occur in quantum field theory. These ideas are basic to all modern theories of elementary particle physics. They are usually analyzed within a Feynman integral framework. *See* ELEMENTARY PARTICLE; NONRELATIVISTIC QUANTUM THEORY; PHASE TRANSITIONS; QUANTUM FIELD THEORY; STATISTICAL MECHANICS; SYMMETRY LAWS (PHYSICS). Barry Simon

Bibliography. R. Cenni et al., *Feynman Integrals in Theoretical, Nuclear, and Statistical Physics*, 1989; R. P. Feynman, Space-time approach to nonrelativistic quantum mechanics, *Rev. Mod. Phys.*, 20:367–387, 1948; R. P. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals*, 1965; M. Kac, On some connections between probability theory and differential equations, *Proceedings of the 2d Berkeley Symposium on Mathematical Statistics Problems*, pp. 189–215, 1950; B. Simon, *Functional Integration and Quantum Physics*, 1979.

Fiber bundle

A decomposition of a space E into a family of identical subspaces (fibers). The space of subspaces is called the base space. Fiber bundles arise naturally in many physical situations. The theory of fiber bundles has been applied to gauge theory in physics, and there is a lively interaction centered on ideas related to fiber bundles and gauge theories in which significant results have been contributed in both areas. The development of the mathematical theory of fiber bundles was begun in the 1930s and has numerous applications within mathematics.

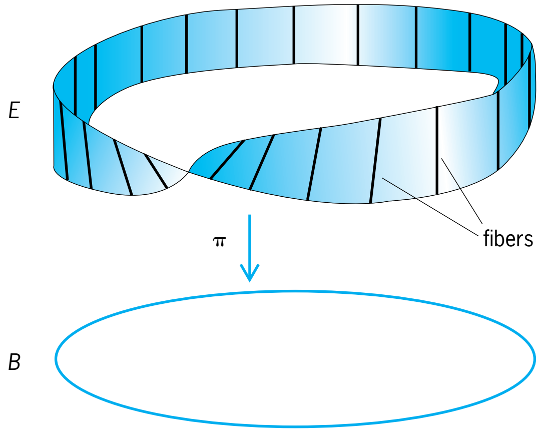
Definition. A fiber bundle consists of three spaces, a fiber F, a total space E, and a base B, together with a map $\pi\colon E \to B$ from the total space to the base. These sets and the map π are supposed to satisfy the following requirements:

1. The spaces F, E, and B are topological spaces: A topology is specified by giving meaning to the phrase: two points are close to one another. It is required that each of the spaces F, E, B has such a notion of closeness. *See* TOPOLOGY.

2. The map π is required to preserve the notion of closeness; if points x and y are sufficiently close to one another in E, then the points $\pi(x)$ and $\pi(y)$ are close in B; that is, π is a continuous map.

3. Each fiber can be identified with F: The map $\pi\colon E \to B$ decomposes E into subsets; two points are in the same subset (or fiber) if they have the same image in B. The set of points in E that map to a given point b in B is called the fiber over b. It is required that each fiber be topologically equivalent (that is, homeomorphic) to F.

4. The total space E has a local product structure: For a set U of points in B that are sufficiently close together, the points $\pi^{-1}(U)$ in E that map to U can be identified with the set of pairs (b, x) with b an



Möbius strip, the simplest nontrivial example of a fiber bundle.

element of U and x an element of F; that is, $\pi^{-1}(U)$ is homeomorphic to $U \times F$.

Examples. The trivial examples of fiber bundles are those where E is a product space $E = B \times F$, the set of pairs (b, x), where b is a member of B, and x is a member of F.

Möbius strip. The simplest nontrivial example of a fiber bundle is the Möbius strip (see **illus.**). It is constructed by twisting one end of a piece of paper and then gluing the two ends together. The fibers are line segments, and the space of line segments is a circle. The map π takes all the points in a fiber into the same point on the circle.

Tangent bundles. Tangent bundles are among the earliest examples of fiber bundles. For spaces M (called manifolds) which are locally like a euclidean space, the space $T(M)$ consisting of pairs (x, v) with x a point in M and v a vector tangent at x to a path in M is called the tangent bundle to M; the fibers form a vector space, and the base space is M. Fiber bundles in which the fibers are vector spaces are called vector bundles. *See* MANIFOLD (MATHEMATICS).

Applications to physics. Laws of physics are often invariant under the action of an appropriate group. Fiber bundles arise in physics in situations where such a group appears as an internal or local symmetry group; the local symmetry group is the fiber. *See* GROUP THEORY; SYMMETRY LAWS (PHYSICS).

For example, by Maxwell's equations, a vector potential A for a magnetic monopole of charge g located at the origin in three-dimensional euclidean space, $\boldsymbol{R}^3$, would have a curl whose magnitude is equal to g/r^2, where r is the distance from the origin to the point p at which the curl of A is evaluated, and whose direction is that of a vector directed from the origin to p. There is no vector potential that both satisfies this condition and is defined on all of $\boldsymbol{R}^3 - (0, 0, 0)$ [three-dimensional space with the origin removed]. However, potentials that satisfy this condition can be defined on sufficiently small open sets in $\boldsymbol{R}^3 - (0, 0, 0)$. For two such open sets, a gauge transformation, Ψ is a mapping from the intersection of U and $\tilde{U}$ to $U(1)$, the group of rotations of a circle, here represented by the complex numbers of modulus one. The vector potentials A, defined on U, and $\tilde{A}$, defined on $\tilde{U}$, are said to be related by the gauge transformation Ψ if, at each point in the intersection of U and $\tilde{U}$, A is equal to the sum of $\tilde{A}$ and a term that is equal to $i\hbar c/e$ times the product of the gradient of Ψ with Ψ^{-1}. Here, $\hbar$ is Planck's constant divided by 2π, c is the speed of light, and e is the electric charge of the monopole. Potentials that are related by gauge transformations and defined on sets whose union covers all of $\boldsymbol{R}^3 - (0, 0, 0)$ determine a fiber bundle with fiber $U(1)$ and base $\boldsymbol{R}^3 - (0, 0, 0)$. The condition that the potentials are related by gauge transformations is equivalent to Dirac's quantization condition relating the electric and magnetic charge; namely, that $2ge/(\hbar c)$ is an integer. *See* CALCULUS OF VECTORS; COMPLEX NUMBERS AND COMPLEX VARIABLES; GAUGE THEORY; MAGNETIC MONOPOLES; MAXWELL'S EQUATIONS.

Applications within mathematics. There are numerous, fundamental applications of fiber bundles within mathematics. Many of these applications are obtained through associating numerical invariants to vector bundles. The geometry of manifolds is then studied in terms of invariants of the tangent bundle of the manifold.

One such application involves the concept of cobordism. Two manifolds that are disjoint (have no elements in common) and whose union can be realized as the boundary of another manifold are called cobordant. A manifold that can be realized as the boundary of manifold is called cobordant to zero. For example, a circle is the boundary of disk and hence is cobordant to zero; likewise, a two-dimensional sphere is the boundary of a solid ball. However, the two-dimensional space whose points are lines through the origin in euclidean three-space is not cobordant to zero.

A set of invariants, called Stiefel-Whitney numbers, of vector bundles detect whether a manifold is cobordant to zero, and whether two manifolds are cobordant: Specifically, the condition that a manifold be cobordant to zero is equivalent to the condition that all Stiefel-Whitney numbers of the tangent bundle of the manifold are zero; and the condition that two manifolds are cobordant is equivalent to the condition that the two manifolds have the same Stiefel-Whitney numbers. Richard Porter

Bibliography. D. Husemoller, *Fiber Bundles*, 3d ed., 1993; K. B. Marathe and G. Martucci, *The Mathematical Foundations of Gauge Theories*, 1992; R. Porter, *Introduction to Fibre Bundles*, 1977; N. E. Steenrod, *The Topology of Fibre Bundles*, 1951.

Fiber crops

Many plants are grown as fiber crops because of their content or yield of fibrous material which is used for many commercial purposes and for home industry. The production and use of plant fibers date back many centuries and form the basis for one of the earliest industrial processes. Fibers may be extracted from various parts of different plants. *See* NATURAL FIBER; PLANT ANATOMY.

Long, multiple-celled fibers can be subdivided into hard, or leaf, fibers that traditionally are used for cordage, such as sisal for binder and baler twine and abaca or manila hemp for ropes; soft, or bast (stem), fibers that are used for textiles, for example, flax for linen, hemp for small twines and canvases, and jute and kenaf for industrial textiles such as burlap; and miscellaneous fibers that may come from the roots, such as "broom" root for brushes, or stems, as Spanish moss for upholstery, or fruits, as coir from coconut husks for cordage and floor coverings. *See* FRUIT; LEAF; ROOT (BOTANY); STEM.

There are many exceptions to these uses. For example, some sisal near the sources of production is used for bags, and jute is used for cordage. However, the large quantities that enter international trade are used as indicated above.

Short, one-celled fibers come from the seeds or seed pods of plants such as cotton and kapok. Cotton is the world's most widely grown and used textile fiber. Kapok, because of its low tensile strength, goes primarily into upholstery, pillows, and life preservers. Many plant fibers are being replaced wholly or in part by synthetic or manufactured fibers. Mixtures of natural and synthetic fibers often exhibit qualities superior to either used alone. *See* ABACA; CELLULOSE; COIR; COTTON; FLAX; HEMP; JUTE; KAPOK TREE; LINEN; MANUFACTURED FIBER; RAMIE; SISAL.

Elton G. Nelson

Fiber-optic circuit

The path of information travel, usually from one electrical system to another, in which light acts as the information carrier and is propagated by total internal reflection through a transparent optical waveguide. An electrooptic modulator and an optoelectric demodulator are required to convert the electrical signals into light and back again at the transmit and receive ends of the link, respectively.

Application. A fiber-optic circuit, or link, is used for data transmission when a shielded twisted pair of wires or a coaxial cable fails to meet the performance criteria of the system designer. Depending upon fiber type, the product of distance and bandwidth of a fiber can be tens to thousands of times greater than that for electrical transmission through wire. An optical communication fiber is a nearly perfect waveguide for light, meaning that little or no energy escapes through radiation. Thus, the data traveling in the fiber are secure from eavesdropping, as well as being harmless in or around equipment sensitive to electromagnetic interference. Telecommunication optical fibers have a very small diameter, 5–10 micrometers (0.0002–0.0005 in.), which allows telecommunication cables to be fabricated with a much higher packing density. Furthermore, the most common materials used to make the fibers, silica and plastic, are less dense than copper, making the cable lighter. Lastly, since the fiber is a dielectric, it can be used in volatile or sensitive environments that require electrical isolation. *See* COMMUNICATIONS CABLE; ELECTRICAL INTERFERENCE; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; WAVEGUIDE.

Optical transmitters. In order to pass electrical information through a fiber, an electrooptic modulator, or transmitter, is used to convert electrical impulses into modulated light, which is then launched, or focused, by the transmitter into the fiber (**Fig. 1**). The transmitter generally consists of a silicon integrated circuit that converts input voltage levels from a personal computer, workstation, or telecommunications company mainframe computer into current pulses. These current pulses, in turn, drive a light-emitting diode (LED). The output light from the diode is then focused by one or more lenses into the fiber. Faster and more expensive circuits use laser diodes, which are semiconductor chips that emit coherent laser light at a very low cost. *See* INTEGRATED CIRCUITS; LASER; LIGHT-EMITTING DIODE.

Optical receivers. An optoelectric demodulator, or receiver, is used to convert the detected light back into electrical impulses at the other end of the fiber (Fig. 1). Light from the fiber is focused onto a reverse-biased *pn*-junction photodiode that generates an electron-hole pair for each photon impinging on or near its active area. The applied reverse bias creates an electric field across the *pn* junction, which causes the generated electron-hole pairs to be swept apart and give rise to photocurrent before they can recombine. Systems requiring higher speed or sensitivity use more expensive PIN (positive-intrinsic-negative) diodes. Another circuit, usually a silicon integrated circuit, amplifies this electron-hole current and converts it into voltage levels suitable for interfacing with the computer at the receiving end. *See* PHOTODIODE; PHOTOVOLTAIC CELL.

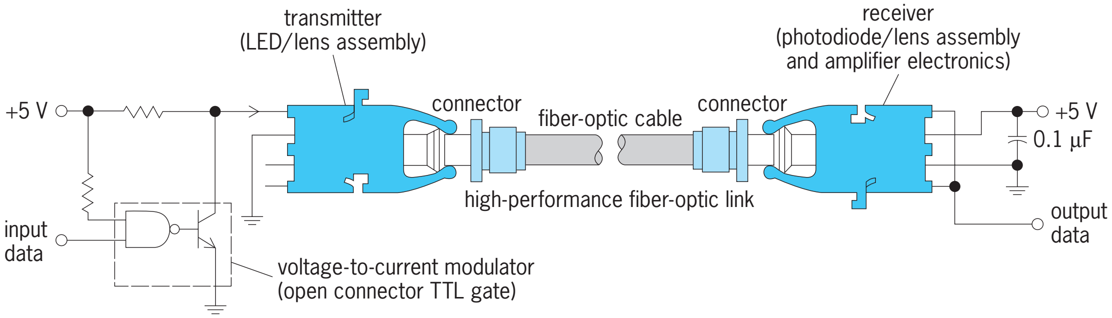


Fig. 1. Typical interface circuit for fiber-optic link that transmits data at 5 megabits per second. (*Hewlett-Packard*)

Link budget. The parameters involved in the design of a fiber-optic link include the optical output power, spectrum, fiber loss, dispersion, extinction ratio, bandwidth, rise and fall times, and near- and far-field radiation patterns of the transmitter; and the sensitivity, bandwidth, signal-to-noise ratio, bit error rate, eye pattern, and intersymbol interference of the receiver. Link design involves trade-offs with these parameters and the costs to obtain the desired link budget. The link budget is the most important parameter in the design of the system. It is defined as the power detected at the receiver in excess of that required to maintain a given signal-to-noise ratio while the link is operating at the maximum specified frequency; it is expressed in decibels (dB). The link budget makes it possible to calculate how far the link will carry signals without a repeater in attenuation-limited systems, or how many connectors and splices can be used at a given distance in dispersion-limited links. *See* SIGNAL-TO-NOISE RATIO.

Links. The simplest and least expensive link uses plastic fiber and operates with visible light having a typical peak wavelength of about 665 nanometers. The transmitter is usually a transistor-transistor logic (TTL) line driver which shunts current from a resistor and an inexpensive aluminum gallium arsenide (AlGaAs) light-emitting diode, while the receiver uses a silicon photodiode integrated with the receiver electronics. Links of this type generally transmit data at 5 megabits per second and are limited to several tens of meters because of the high optical attenuation of the plastic cable, typically 0.25 dB/m or 0.075 dB/ft (**Fig. 2**). *See* LOGIC CIRCUITS.

To transmit data over longer distances or at higher rates, silica fiber, 820-nm light-emitting diodes, and silicon PIN detectors are used. Silica fiber has a typical attenuation of 2.8 dB/km (4.5 dB/mi) at 820 nm. Silica fiber also has a much higher bandwidth than plastic fiber, allowing data transmission at hundreds of megabits per second. Optical dispersion limits the 820-nm light-emitting diode links to about 100 megabits per second at 1 km (0.6 mi; Fig. 2).

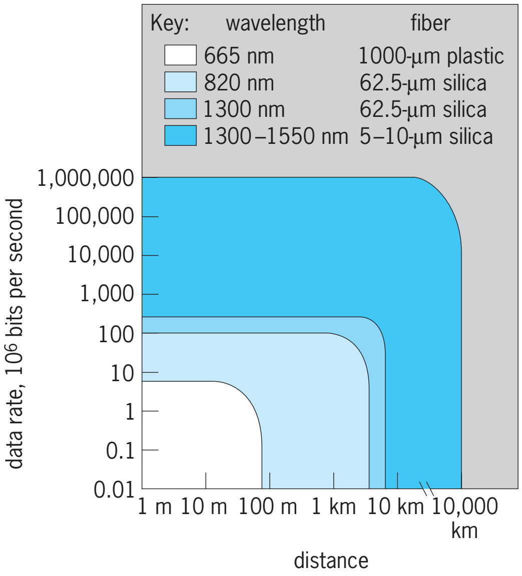


Fig. 2. Performance of four types of fiber-optic links. The limits shown for wavelengths of 1300–1550 nm over 5–10-μm silica fiber are the expected limits for 2010 with wavelength division multiplexing (WDM).

Optical dispersion is near zero with silica fibers at 1300 nm. Light-emitting diodes and PIN photodiodes made of indium phosphide and indium gallium arsenide phosphide (InP/InGaAsP) operate at 1300 nm, have longer operating lifetimes than the AlGaAs light-emitting diodes, and are capable of data rates of over 200 megabits per second at 2 km (Fig. 2). The transmitter and receiver electronics as well as packaging parasitics are much more complicated and interrelated at these data rates. The 820-nm laser diodes can be used for longer distances at higher data rates, but their reliability is much less than that of both 820- and 1300-nm light-emitting diodes. Typical systems such as those described above operate over several kilometers at data rates of several hundred megabits per second. By replacing light-emitting diodes with vertical cavity surface emitting lasers (VCSELs), these local area network links are capable of operating in the gigabit-per-second region. *See* LOCAL-AREA NETWORKS.

Commercial telecommunication links operate at the rate of low tens of gigabits per second. They use single-mode fiber with laser transmitters operating between 1300 and 1580 nm. Trade-offs between modal and chromatic dispersion versus cost, wavelength, fiber diameter, data rate, and distances from tens to hundreds of kilometers without repeaters are made, depending on what is most important for the desired link (Fig. 2). Most telecommunication transmitters use 1530–1580-nm edge emitting semiconductor lasers to be compatible with optical amplifiers.

Optical amplifiers. Despite the tremendous information-carrying capacity of optical fibers, the performance of most communication systems is limited by electronics speed in signal regeneration. Erbium-doped fiber amplifiers (EDFAs) overcome the speed and cost limitations of electronically "repeatered" communication systems and have the advantages of temperature stability, quantum-limited noise figure, polarization insensitivity, and immunity to interchannel crosstalk. Erbium-doped fiber amplifiers expand on the physics of laser diodes allowing amplification of light. For each photon that goes into an erbium-doped fiber amplifier, many coherent photons come out, thus providing amplification with minimal signal degradation. Erbium-doped fiber amplifiers provide a more efficient and cost-effective technology for both transoceanic and terrestrial communication links where the demands for bandwidth continue to grow at an exponential rate as telecommunication markets grow and the Internet gains popularity. *See* INTERNET.

Optical routers. With the surge of networking in the 1990s, optical routers are an obvious extension. Networks allow multiple computers to connect all at once. A router takes data from one source computer link and directs it to the proper data link going toward the destination computer on a network. Typical routers are packed with fiber-optic transceivers where input optical data are converted to electricity, routed to the proper transmitter, and then converted back to light and sent to the next router or destination computer. This optical-to-electrical-to-optical conversion is both relatively expensive and slow. Advanced routers now are doing this without the conversion and are replacing the electronics with micromachined mirrors that steer the light. These routers direct data traffic from one segment of a network to another at up to 10 terabits, or 10^{13} bits, of data per second. These advanced routers can direct the information equivalent of whole libraries in seconds.

Wavelength-division multiplexing. Wavelength-division multiplexing (WDM) is another means of getting more data though a fiber, thus delaying or eliminating the need to install more strands of fiber. Wavelength-division multiplexing is accomplished by encoding data in different wavelengths of light and multiplexing these in a single-mode fiber. These separate single-mode fibers are then fused together and spliced to a single fiber that is usually already installed. At the other end of the fiber, the light is similarly demultiplexed and converted to electrical signals with wavelength-tuned receivers. Usually, the detectors are identical, and narrow Bragg grating filters are used to select which wavelength will hit each detector. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Kirk W. Lindstrom

Bibliography. P. C. Becker et al., *Erbium-Doped Fiber Amplifiers: Fundamentals and Technology*, 1999; D. Derickson (ed.), *Fiber Optic Test and Measurement*, 1997; I. P. Kaminow and T. L. Koch (eds.), *Optical Fiber Telecommunications III and IIIb*, 1997; H. Kressel (ed.), *Semiconductor Devices for Optical Communication*, 2d ed., updated 1990; J. C. Palais, *Fiber Optic Communication*, 4th ed., 1998.

Fiber-optic sensor

A sensor that uses thin optical fibers to carry light to and from a location to be probed. In performing the sensing, light can be lost from the fibers or modified in velocity by the action of the phenomena on the fiber. Fiber-optic sensors are ideal for probing in remote or hostile locations, where miniature sensors are required such as in the body, or where extreme sensitivity is required. Two classes of fiber sensors have evolved: intensity sensors, in which the amplitude of light in the fiber is changed during sensing, and interferometric sensors, in which the velocity of light or its phase is modified during sensing. The latter class has proved to be extremely sensitive; intensity sensors are used where moderate performance is acceptable and lower cost is important.

Intensity sensors. The intensity sensor configurations shown in **Fig. 1** illustrate several techniques used in fiber sensors to probe the environment; over 100 sensor designs are currently available. Depending on their design, fiber sensors can respond to pressure, temperature, liquid level, position, flow, smoke, displacement, electric and magnetic fields, chemical composition, and numerous other conditions.

With the pressure sensor (Fig. 1*a*), the displacement of a diaphragm which responds to pressure is detected by light probes. Light is carried to the diaphragm by the central fiber and is reflected. The amount of light captured by the fibers surrounding the central fiber is proportional to the displacement of the diaphragm and thus to the pressure. *See* PRESSURE MEASUREMENT.

In the liquid-level sensor (Fig. 1*b*), light carried by the fiber is reflected up when the fiber is in air; however, in a liquid the reflection is dispersed and light lost into the liquid. By monitoring of the reflected light at the input end, the presence or absence of the liquid can be determined. *See* LEVEL MEASUREMENT.

The remote-detection sensor (Fig. 1*c*) uses the fiber to carry light to the location to be probed. Light is scattered by the gas or liquid at the end of the fiber and partially recaptured by the fiber. Analysis of the scattered light at the input end can yield information on chemical composition or the presence of species such as pollutants in water. A sensor of this type is used on over 1000 ships to monitor oil contamination in discharge water. These sensors have also been tested clinically to determine pH and oxygen and carbon dioxide pressure in blood. The small size of the fiber probe (100 micrometers) makes these sensors ideal for invasive testing.

Encoder sensors (Fig. 1*d*) utilize a bundle of fibers to read bar codes printed on the sample. A wide variety of sensitized-tip sensors (Fig. 1*e*) for temperature, pressure, and other parameters utilize a material attached to the fiber with optical properties that change with the value of the measured quantity. Temperature sensors using this principle typically have sensitivities of 0.1°C (0.2°F) with ranges over 200°C (360°F). In the microbend sensor (Fig. 1*f*), light can be lost from the fiber by squeezing it at sharp,

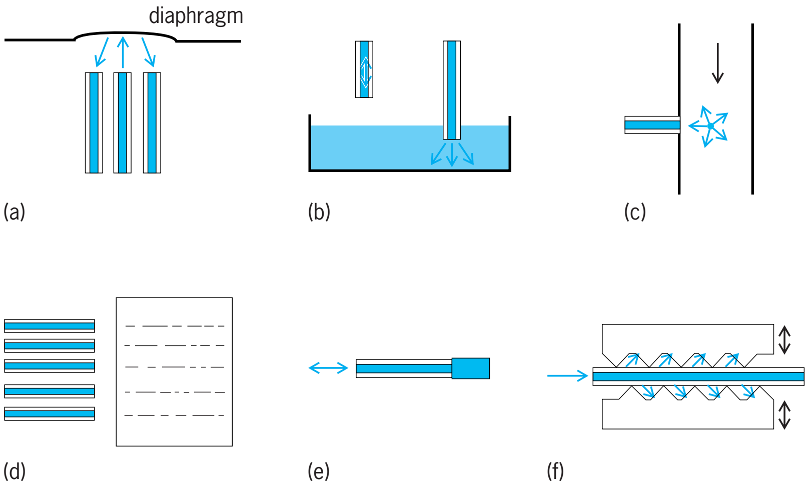


Fig. 1. Intensity fiber-optic sensors. (*a*) Pressure. (*b*) Liquid-level. (*c*) Remote detection. (*d*) Encoder. (*e*) Sensitized-tip. (*f*) Microbend.

distinct points. Pressure, temperature, liquid-level, and intruder sensors have all been built by using this principle.

Interferometric sensors. Optical interferometry is one of the most sensitive means of detecting displacements as small as 10^{-13} m. Interferometric fiber sensors apply this technology to sense many physical phenomena.

Fiber gyro. A fiber gyro based on the Sagnac effect (**Fig. 2***a*) is formed by making a fiber loop which, when rotated, causes the light traveling in both directions in the loop to experience different velocities with or against the rotation. Light is coupled into the fiber from a light source, and the fiber coupler divides the input light equally in both directions. After traversing the loop, the light is again combined in the coupler and detected photoelectrically. The phase shift of the light, which is a measure of the rotation rate, is given by Eq. (1), where λ is the wavelength of

$$\phi_s = \frac{\pi L D \Omega}{\lambda c} \qquad (1)$$

the light, c is the speed of light, L the length of fiber, D the diameter of the loop, and Ω the rotation rate to be measured. By increasing the length of fiber L or the loop diameter D, the gyro can be made more sensitive. Minimum rotation rates of $10^{-4\circ}$/h have been demonstrated in the laboratory. The type of fiber is known as polarization-preserving (Fig. 2*b*); that is, the polarization of the light in the fiber is controlled. Polarization control is required in order to obtain the maximum sensitivity from Sagnac interferometers. *See* GYROSCOPE; POLARIZED LIGHT.

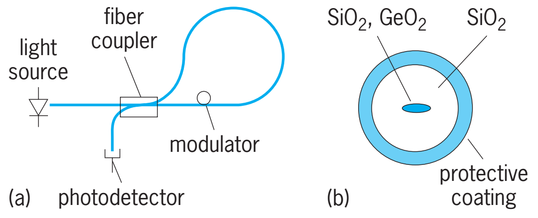


Fig. 2. Fiber gyro based on the Sagnac effect. (*a*) Configuration. (*b*) Cross section of polarization-preserving fiber.

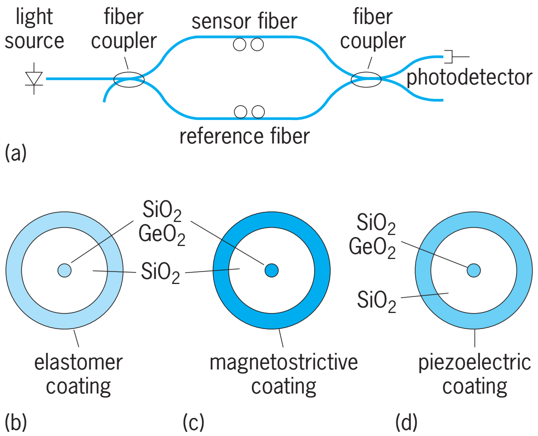


Fig. 3. Interferometric sensor based on Mach-Zehnder interferometer. (*a*) Configuration. (*b*) Cross section of fiber used in acoustic sensors, (*c*) magnetic sensors, and (*d*) electric sensors.

Mach-Zehnder sensor. A second type of interferometric sensor is constructed by using the Mach-Zehnder interferometer (**Fig. 3***a*). Light from a semiconductor laser is divided equally by a fiber coupler into the sensor and reference arms of the interferometer. The sensor fiber is treated with material which responds selectively to the phenomena to be sensed, whereas the reference fiber is selected not to respond. Coatings are applied to the fibers to achieve the desired sensitivity: elastomers for acoustic sensors (Fig. 2*b*), magnetostrictive material for magnetic sensors (Fig. 2*c*), piezoelectric material for electric sensors (Fig. 2*d*), and so on. These coatings stretch the fiber in response to the phenomena being measured; the induced phase shift caused by the stretching is given by Eq. (2), where λ and L are the wave-

$$\phi = \frac{2\pi}{\lambda}\left(\frac{1}{L}\frac{dL}{dP}\right)PL \qquad (2)$$

length of the light and length of fiber respectively and P is the field strength being measured. This phase shift is converted into an amplitude modulation by the passage of the light through the second coupler. The modulated light or signal is detected photoelectrically. Sensitivities equal to or surpassing the best conventional technologies have been achieved in these sensors. For example, fiber hydrophones have demonstrated sensitivities of 30 dB re 1 micropascal (that is, the ability to detect signals whose intensity is 30 dB with respect to a sound wave whose acoustic pressure is 1 μPa) whereas the best piezoelectric acoustic sensors have shown sensitivities of 40 dB re 1 μPa. at 100 Hz; thermal fiber sensors have shown sensitivities of $10^{-7\circ}$C with 5-kHz bandwidth, whereas the best conventional thermometer sensitivities are about $10^{-5\circ}$C. *See* DECIBEL; INTERFEROMETRY; MAGNETOSTRICTION; PIEZOELECTRICITY.

Thomas G. Giallorenzi

Bibliography. K. Grattan and B. Meggitt, *Optical Fibre Sensor Technology*, 1992; B. P. Pal, *Fundamentals of Fibre Optics in Telecommunication and Sensor Systems*, 1993.

Fiber-optics imaging

The use of fiber optics in image transmission, based on the ability of a precisely aligned bundle of optical fibers to transmit an image from one end of the bundle to the other. Each fiber transmits one element of the image so that the complete image is made up of a matrix of dots (pixels) which blend into a recognizable image like a halftone picture on a printed page.

Light is transmitted through individual fibers by means of total internal reflection from the fiber walls. For efficient transmission, each fiber has a highly transparent core (usually glass) coated with a layer (cladding) of lower refractive index (also usually glass). This cladding prevents light from leaking into neighboring fibers (crosstalk) and protects the core from contamination and wear. Fiber bundles may be either flexible or rigid. *See* CROSSTALK; OPTICAL MATERIALS; REFRACTION OF WAVES.

Flexible fiber bundles. Most flexible bundles are made by either winding or leaching. In the former a single or a multiple fiber is drawn from a preform, which is slowly lowered through an electric furnace, and is precisely wound onto a rotating drum or slotted frame. The helical ribbons thus formed are cemented along a short portion of their circumference, stactional area, and permanently bonded at the cemented portion. The loop formed is cut open and the end faces ground and polished to form a coherent image-transmitting fiber bundle.

In the leaching process the preform consists of a core rod with a double-glass cladding, the outer cladding being acid-soluble. This is drawn into small rods 0.04–0.08 in. (1–2 mm) in diameter, which in turn are assembled into a multifiber preform and drawn to a multifiber bundle. In some cases, this process is repeated. The final rigid rod is cut to the desired length and leached (except at the ends) to form a flexible bundle. In both cases, the fibers must be precisely aligned to avoid image distortion and loss of resolution.

A third (hybrid) form of the flexible image bundle consists of a fused bundle of clad silica-based fibers drawn in a silica tube to a diameter that is small enough to be flexible. In this all-fused type, the degree of flexibility strongly depends on the bundle diameter, which is not the case in the leached or wound types of bundles.

The effective image resolution is determined by the total number of fibers (pixels) which the bundle contains. The number is limited by the size of the bundle, since the minimum size of an individual fiber is about 8 micrometers (5 μm in all-fused types). In general, wound bundles are used for larger sizes from about 0.12 in. (3 mm) to as much as 1 in. (25 mm) in diameter. The leached type ranges from 0.12 in. (3 mm) to about 0.02 in. (0.5 mm). The total number of pixels can be as low as 2000 to more than a million in the largest wound bundles.

Flexible image bundles are used in a wide variety of industrial and medical fiberscopes or endoscopes to transmit the image from an objective lens at the distal end to an eyepiece in the control handle. In general, several components compete for space within the flexible shaft of the instrument, including the control wires, the fiber light guides (to illuminate the object), and channels to admit instruments.

Fiberscopes range in diameter from 0.04 in. (1 mm) to 0.6 in. (15 mm) or more. The smaller versions are of necessity simpler with little or no control of distal articulation and at most a single operating channel. Larger-diameter instruments have remote-tip articulation either in a single plane or, with dual control knobs, in all planes. All are equipped with fiber-optic illumination.

Industrial fiberscopes ranging in length from 3 to 8 ft (1 to 2.5 m) are used in aircraft engine inspection, for example, examination of turbine blades. Longer industrial fiberscopes are used for pipe and weld inspection, including those in nuclear power plants. Here radiation-resistant silica fiber bundles are employed for image transmission and illumination.

Medical endoscopes range in size from less than 0.08 in. (2 mm) for viewing inside the arteries (cardioscope) to 0.6 in. (15 mm) for examining the colon (colonoscope). The latter has an operating channel large enough to perform surgery via remotely controlled forceps or electrosurgical snares. Intermediate-sized endoscopes are used to view the bronchi (bronchoscope), the kidneys (nephroscope), the bladder (cystoscope), the throat (laryngoscope), and the stomach (gastroscope).

Rigid or fused fiber bundles. These are made from glass preforms by a series of drawing steps leading to multifibers and multi-multifibers, which are fused under controlled heat and pressure into a block or boule of fibers. Precise alignment is achieved by assembling multi-multifibers that are large enough to be accurately fitted together. The resulting boule contains millions of individual fibers with sizes ranging from 10 to 4 μm.

Since many applications of fused-fiber components involve the transmission of diffused light or light from large-aperture lenses, the components are designed to provide a wide acceptance angle (high numerical aperture). For this, glasses for the core and cladding are selected with a large refractive index difference. *See* GEOMETRICAL OPTICS.

Fused-fiber bundles often contain means for eliminating stray light caused by light entering the cladding and scattering throughout the component. This is done by adding a pattern of minute light-absorbing glass fibers among the transmitting fibers. These intercept the stray light but do not absorb the signal light.

Fused-fiber boules are the starting point for various image-transmitting components, including faceplates (windows) for image-intensifier tubes and cathode-ray tubes, image inverters (twisters), magnifiers, and rigid conduits. Faceplates of various shapes and sizes are cut and machined directly from the boule. The entrance and exit faces are ground and polished by normal lens-making processes. *See* CATHODE-RAY TUBE; IMAGE TUBE (ASTRONOMY); LIGHT AMPLIFIER.

One faceplate of an image intensifier transfers light from the image formed on its entrance face to the photocathode layer deposited on its exit face. This inner face may be concave to correct the aberrations of the electron lens in the intensifier. A second faceplate transmits the image formed on a phosphorescent screen deposited on the inner face of this plate to the exit face, where it is coupled to another fiber-optic component or an eyepiece.

Image intensifier tubes are used in military night-vision devices, low-light television cameras, and astronomical telescopes. In a cathode-ray tube the fiber-optic faceplate is used to transmit the image on the phosphor screen to the exit face for printing directly onto photographic material.

Image inverters and magnifiers require further hot-forming operations in their manufacture. In the former the boule is clamped at both ends, heated in the middle, and one end twisted through 180° (or any desired angle). Fiber-optic magnifiers are similarly made by stretching instead of twisting the boule. The resulting hourglass bundle is cut into two tapers and the ends are polished. Since the diameters of the two ends of the taper are different, they can be used to magnify or reduce an image. Image conduits are made by stretching (or drawing) a boule into a long rod of uniform diameter. Although rigid, these can be bent by heating into various shapes to pass around corners. *See* OPTICAL FIBERS.

Walter P. Siegmund

Fibrinogen

The major clot-forming substrate in the blood plasma of vertebrates. Though fibrinogen represents a small fraction of plasma proteins (normal human plasma has a fibrinogen content of 2–4 mg/ml of a total of 70 mg protein/ml), its conversion to fibrin causes a dramatic gelation which blocks the flow of blood. Upon injury, sufficient amounts of the clotting enzyme, thrombin, are generated in about 5 min clotting time to produce a gel.

Although clotting in the circulation (thrombosis) can be extremely dangerous, clotting is an essential and normal response (hemostasis) for preventing the loss of blood. Individuals born with the hereditary absence of fibrinogen (afibrinogenemia) suffer from severe bleeding, which can be counteracted by transfusing normal plasma or purified fibrinogen. There are also rare cases of abnormal fibrinogens (dysfibrinogenemia) which may be associated with a tendency for bleeding. Depletion of fibrinogen might occur by disseminated intravascular coagulation when substances which promote clotting enter the circulation (as in placenta abruptio). Premature degradation of fibrinogen by the proteolytic enzyme plasmin could also lower the fibrinogen concentration of plasma to levels where life-threatening hemorrhage might ensue. Malignant cells produce large amounts of activators of this fibrinolytic system, and the problem is often compounded by the fact that the fibrinogen-degradation products themselves interfere with clotting. Drugs such as ϵ-aminocaproic acid, which inhibit the conversion of plasminogen to plasmin, might be useful in such a situation. *See* HEMORRHAGE; THROMBOSIS.

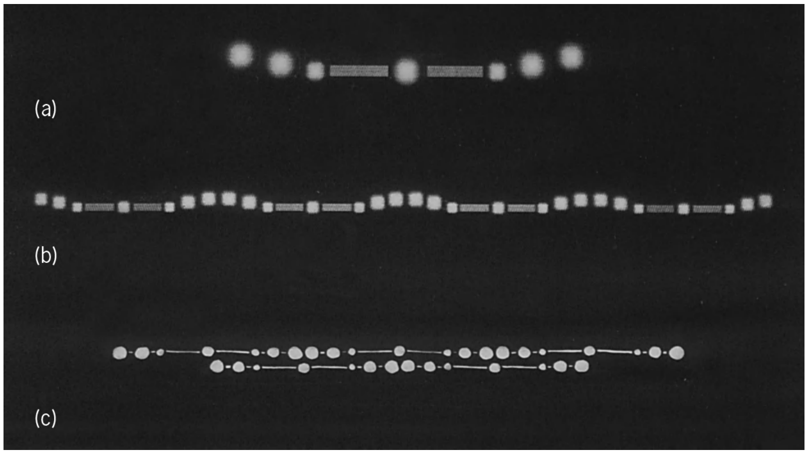


Fig. 1. Fibrinogen models. (*a*) Computer image of the symmetrical molecule. (*b*) Model for end-to-end aggregation. (*c*) Half-staggered molecular overlap between side-to-side aggregated filaments in a clot network. (*Parts a and b from C. Cohen et al., The structure of fibrinogen and fibrin, pt. I, Ann. N. Y. Acad. Sci., 408:194–213, 1983;* c *from J. S. Weisel, G. N. Phillips, Jr., and C. Cohen, pt. II, Ann. N. Y. Acad. Sci., 408:367–378, 1983*)

Fibrinogen is synthesized by the hepatocytes in the liver, and the synthetic rate can be stimulated by hormones. Significant amounts of carbohydrates become attached to the protein before it is secreted into the circulation; fibrinogen is a plasma glycoprotein. Alterations in its carbohydrate composition as found in some liver diseases can give rise to abnormal fibrinogens with defective clotting properties.

Structure. In electron micrographs, fibrinogen has a length of about 45 nanometers, and the molecule appears to be symmetrical. The structure can be represented best by a model comprising globular as well as rodlike regions (**Fig. 1***a*). The diameters of the nodular domains (starting from either end) are 4.8, 4.8, 3.4, and 4.5 nm. The rods are 2.2 nm thick. Though not shown in Fig. 1, there are also two flexible appendages (a short and a long one) protruding from both ends of the molecule. The longer appendages are highly exposed, and they can be easily removed by proteolytic enzymes.

During clotting, an end-to-end association ensues (Fig. 1*b*), followed by a side-to-side aggregation of the filaments (Fig. 1*c*), producing a half-staggered overlap of individual molecules with a characteristic cross-striation of about 22.5 nm. In this array, the central nodule of one molecule interacts with the two end-to-end associated, terminal globules in the neighboring filament.

Composition. As far as chemical composition is concerned, human fibrinogen, with a molecular weight of 337,897, comprises two sets of three open polypeptide chains (Aα, Bβ, and γ) held together by disulfide bonds; thus, the composition can be expressed as $A\alpha_2B\beta_2\gamma_2$. Only the B$\beta$ and γ chains carry carbohydrate residues. Each of the three polypeptides is encoded by a distinct messenger ribonucleic acid, which indicates that the synthesis of fibrinogen is the result of the concerted expression of three genes. The Aα chain is the largest (mol wt 66,026), followed by Bβ (mol wt 54,358) and γ chain (mol wt 48,529). Though the exact mode of assembly of these constituent chains in the model given in Fig 1 is not yet known, they are arranged such that all N termini are located in the central globular region and all C termini are found on the outermost structures. The long, flexible appendages protruding from each end of the molecule comprise the C-terminal regions of the Aα chains and the short appendages the C termini of γ chains. The complete amino acid sequences of all chains of human fibrinogen are known. Considerable evolutionary variations exist in vertebrates

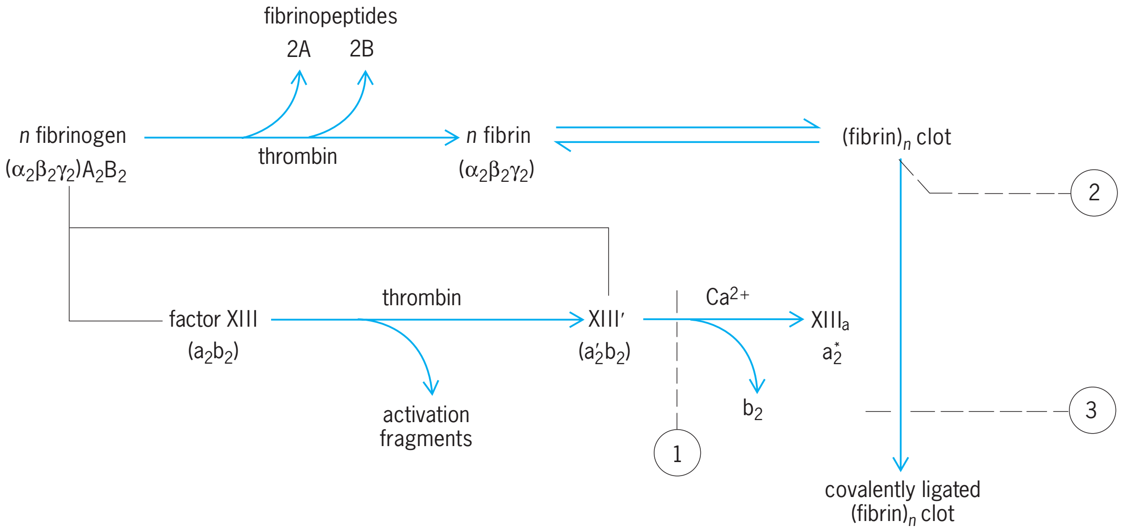


Fig. 2. Physiological pathway for the clotting of fibrinogen in normal plasma. Clotting time is measured by forming the $(\text{fibrin})_n$ network; this, however, is not the end product. Unless this clot undergoes ligation by γ-glutamyl-ϵ-lysine cross-links, bleeding may ensue. Such disorders of fibrin stabilization occur in the hereditary absence of factor XIII or when inhibitors to one of the reaction steps (marked 1–3) appear in the blood circulation. (*After T. H. Spaet, ed., Progress in Hemostasis and Thrombosis, vol. 5, pp. 245–290, Grune and Stratton, 1980*)

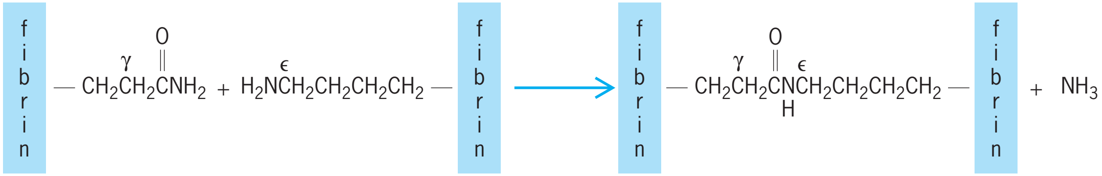

$$\text{fibrin}-\overset{\gamma}{CH_2}CH_2\overset{O}{\overset{\|}{C}}NH_2 + H_2\overset{\epsilon}{N}CH_2CH_2CH_2CH_2-\text{fibrin} \longrightarrow \text{fibrin}-\overset{\gamma}{CH_2}CH_2\overset{O}{\overset{\|}{C}}\underset{H}{\overset{\epsilon}{N}}CH_2CH_2CH_2CH_2-\text{fibrin} + NH_3$$

Fig. 3. Formation of γ-glutamyl-ϵ-lysine peptide bridges between fibrin molecules is the last enzyme-catalyzed reaction in blood clotting.

with regard to the sequences of N-terminal peptides in the Aα and Bβ chains (called fibrinopeptides A and B) and in the exposed C-terminal half of the Aα chain.

Clotting mechanism. It is now possible to reconstruct with reasonable certainty the sequence of events surrounding the clotting of fibrinogen in normal plasma, and each of the reaction steps in the scheme in **Fig. 2** can be studied separately with the use of purified components. Clotting is regulated by two enzymes, thrombin and factor $XIII_a$ (fibrinoligase, activated fibrin-stabilizing factor, transglutaminase). Thrombin is a protease similar to trypsin but of much greater specificity. It exerts a dual control by regulating the rate of fibrin formation as well as producing factor $XIII_a$ by processes of limited proteolysis. In the fibrinogen-fibrin conversion, about 97% of the parent protein becomes fibrin and 3% appears as fibrinopeptides A and B; in the reaction of thrombin with factor XIII, activation fragments are released from the N termini of the *a* subunits of factor XIII. The presence of fibrinogen greatly accelerates the latter reaction so that the production of fibrin and the activation of factor XIII are harmonized. In the plasma milieu, the fibrin molecules readily aggregate into a clot. In order to obtain a clot structure of a strength sufficient to stem bleeding, it is necessary for the thrombin-modified factor XIII (indicated as XIII′ in Fig. 2) to be activated to $XIII_a$. This is promoted by fibrinogen and is accomplished by the action of Ca^{2+} ions which dissociate the catalytic subunits (a^*_2) from the inert, carrier subunits (b_2). Factor $XIII_a$ acts as a transamidating enzyme which strengthens the fibrin clot by creating a few γ-glutamyl-ϵ-lysine cross-links (**Fig. 3**). Without such cross-links, a clot structure would be like a brick wall without mortar.

Individuals with the hereditary absence of factor XIII often suffer from severe bleeding, even though their clotting times are in the normal range. In addition, various molecular disorders of fibrin stabilization are known in which some acquired inhibitor suddenly makes its appearance in the circulation, directed either against the activation of factor XIII or against the cross-linking of fibrin by $XIII_a$. Several such inhibitors were found to be autoimmune IgG antibodies. Primary amines also specifically inhibit cross-linking without interfering with fibrin formation. *See* BLOOD; HEMOPHILIA; IMMUNOGLOBULIN.

L. Lorand

Bibliography. D. W. Chung et al., Cloning of fibrinogen genes and their cDNA, *Ann. N.Y. Acad. Sci.*, 408:449–456, 1983; R. F. Ebert, *Index of Variant Human Fibrinogens*, 1991; C. Y. Liu and S. Chen (eds.), *Fibrinogen, Thrombosis, Coagulation, and Fibrinolysis*, 1990; L. Lorand, Fibrinopeptide: New aspects of the fibrinogen-fibrin transformation, *Nature*, 167:992, 1951.

Fibromyalgia syndrome

A chronic condition characterized by widespread pain in muscles, tendons, and ligaments, and the presence of multiple tender points on the body. The illness occurs in up to 5% of the general population, and approximately 90% of affected individuals are women.

Symptoms and diagnosis. Like many other syndromes, fibromyalgia has no single specific feature but is a symptom complex of self-reported or elicited findings. In 1990, the American College of Rheumatology published diagnostic criteria, which include widespread pain for more than 3 months and muscle tenderness in at least 11 out of 18 so-called tender points, places on the body that are sensitive to the application of pressure. Most tender points are located at tendon insertion areas. Besides musculoskeletal pain and mechanical tenderness, most patients with fibromyalgia also complain of insomnia, fatigue, and distress. In addition, fibromyalgia appears to coaggregate with major mood disorders. *See* AFFECTIVE DISORDERS; MUSCLE; MUSCULAR SYSTEM.

Pathogenesis. The hypersensitivity of patients with fibromyalgia is not limited to pain but also includes light, sound, and smell. The cause for this heightened sensitivity is unknown, but central nervous system sensory processing abnormalities have been reported in several studies. Most of these studies found evidence of central sensitization of neurons of the spinal cord and the brain, including areas that are related to pain (such as the dorsal horn of the spinal cord and the thalamus). The pathogenesis of central sensitization in fibromyalgia is unclear but may be related to prior stressors, including infections and traumas. Because most patients report feeling pain in deep tissues, particularly muscles, these structures may play an important role in the initiation and maintenance of central sensitization and pain. No consistent tissue abnormalities have been reported in patients with fibromyalgia that would explain persistent pain. However, focal muscle areas, like trigger points, may be relevant for the pain of fibromyalgia because little nociceptive input from peripheral tissues is required for the maintenance of central sensitization. *See* NERVOUS SYSTEM (VERTEBRATE); PAIN.

Treatment. Treatment of fibromyalgia needs to be individually tailored to each patient's needs. The identification of pain generators is essential for an effective treatment plan. It is also important to assess biopsychosocial abnormalities (particularly anxiety and depression), which can be detected in most patients with fibromyalgia. Pharmacotherapy has been most successful with the use of antidepressant, muscle relaxant, and anticonvulsant drugs. These drugs affect the release of various neurochemicals (such as serotonin, norepinephrine, and substance P) that have a broad range of activities in the brain and spinal cord, including modulation of pain sensation and tolerance. None of these drugs, however, is currently approved by the U.S. Food and Drug Administration for the treatment of fibromyalgia. *See* NORADRENERGIC SYSTEM; PSYCHOPHARMACOLOGY; SEROTONIN.

The efficacy of patients to manage their pain seems to correlate with their functional status (change to physical function). Brain imaging and psychological profiles have identified at least three fibromyalgia subgroups, that is, patients who display high dysfunction, interpersonal distress, or effective coping mechanisms. Such studies provide an explanation of why some treatments seem to be differentially effective in individual patients.

Most patients' symptoms respond to low-dose tricyclic medications, such as amitriptyline and cyclobenzaprine, as well as cardiovascular exercise, cognitive–behavioral therapy, patient education, or a combination of these approaches. Thus, optimal management of fibromyalgia requires a combination of pharmacological and nonpharmacological therapies, with patients and health care professionals using a team approach. *See* PSYCHOTHERAPY.

Roland Staud

Bibliography. D. D. Price et al., Enhanced temporal summation of second pain and its central modulation in fibromyalgia patients, *Pain*, 99:49–59, 2002; R. Staud et al., Abnormal sensitization and temporal summation of second pain (wind-up) in patients with fibromyalgia syndrome, *Pain*, 91(1–2):165–175, March 2001; R. Staud et al., Body pain area and pain-related negative affect predict clinical pain intensity in patients with fibromyalgia, *J. Pain*, 5(6):338–343, August 2004; R. Staud et al., Ratings of experimental pain and pain-related negative affect predict clinical pain in patients with fibromyalgia syndrome, *Pain*, 105(1–2):215–222, September 2003; R. Staud and M. L. Smitherman, Peripheral and central sensitization in fibromyalgia: Pathogenetic role, *Curr. Pain Headache Rep.*, 6:259–266, 2002; C. J. Vierck et al., The effect of maximal exercise on temporal summation of second pain (wind-up) in patients with fibromyalgia syndrome, *J. Pain*, 2(6):334–344, 2001.

Fidelity

The overall input-output accuracy of a signal storage or transmission system. Many attributes contribute to fidelity. In an audio recording-playback system, some of the individual attributes are distortion, bandwidth, signal-to-noise ratio, group delay, and loudspeaker dispersion. Most system designers would agree that noise and distortion are of chief importance. The fidelity of video systems depends on many attributes, including optical quality, bandwidth, dynamic range, and phosphor brilliance, whereas that of a motion picture system depends on optical quality, mechanical stability in camera and projector, and the exposure-density relationships of the film. *See* DISTORTION (ELECTRONIC CIRCUITS); ELECTRICAL NOISE; SIGNAL-TO-NOISE RATIO.

In a given system, fidelity is improved when the linearity of each constituent element in the system has been improved. However, fidelity may be increased by redefining the system. For example, stereophonic sound transmission is inherently of superior fidelity compared with monophonic transmission in that it provides auditory perspective. It can be argued that two stereophonic channels of moderate quality offer greater fidelity than a single monophonic channel of the highest quality, thus underscoring the subjective basis for assessing fidelity. Similarly, most viewers would agree that color video transmission presents greater fidelity than monochrome transmission. *See* CINEMATOGRAPHY; LINEARITY; SOUND-REPRODUCING SYSTEMS; STEREOPHONIC SOUND; TELEVISION. John Eargle

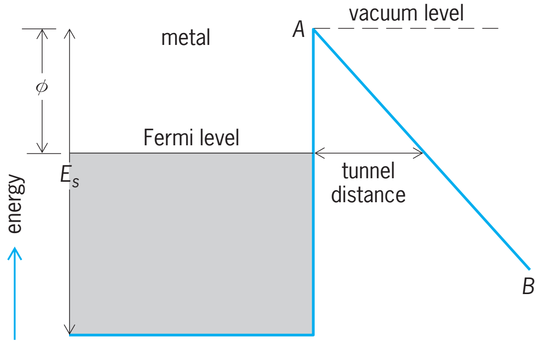
Field emission

The emission of electrons from a metal or semiconductor into vacuum (or a dielectric) under the influence of a strong electric field. In field emission, electrons tunnel through a potential barrier, rather than escaping over it as in thermionic emission or photoemission. The effect is purely quantum-mechanical, with no classical analog. It occurs because the wave function of an electron does not vanish at the classical turning point, but decays exponentially into the barrier (where the electron's total energy is less than the potential energy). Thus there is a finite probability that the electron will be found on the outside of the barrier. This probability varies as $e^{-cA^{1/2}}$, where c is a constant and A the area under the barrier. *See* NONRELATIVISTIC QUANTUM THEORY; PHOTOEMISSION; THERMIONIC EMISSION.

For a metal at low temperature, the process can be understood in terms of the **illustration**. The metal can be considered a potential box, filled with electrons to the Fermi level, which lies below the vacuum level by several electronvolts. The distance from Fermi to vacuum level is called the work function, ϕ. The vacuum level represents the potential energy of an electron at rest outside the metal, in the absence of an external field. In the presence of a strong field F, the potential outside the metal will be deformed along the line *AB,* so that a triangular barrier is formed, through which electrons can tunnel. Most of the emission will occur from the vicinity of the Fermi level where the barrier is thinnest. Since the electron distribution in the metal is not strongly temperature-dependent, field emission is only weakly temperature-dependent and would occur even at the absolute zero of temperature. The current density J is given by the Fowler-Nordheim equation below, where B is a field-

$$J = BF^2 e^{-6.8\times10^7 \phi^{3/2}/F}$$

independent constant of dimensions A/V^2 (A = amperes or state), ϕ is work function in electronvolts, and F is applied field in V/cm. The factor $\phi^{3/2}/F$ is proportional to the square root of the area under the barrier at the Fermi level. Appreciable emission requires fields of 4–7 × 10^7 V/cm, depending on ϕ. *See* FREE-ELECTRON THEORY OF METALS; WORK FUNCTION (ELECTRONICS).



Energy-level scheme for field emission from a metal at absolute zero temperature.

Field emission is most easily obtained from sharply pointed metal needles whose ends have been smoothed into nearly hemispherical shape by heating. Tip radii r_t equal to or less than 100 nanometers can be obtained in this way; because of its small size, the emitter is generally a single crystal. If an emitter is surrounded by a hemispherical anode raised to a voltage V, then the field F is approximately $V/5r_t$ at the emitter. Thus, modest voltages suffice for emission. The electric lines of force diverge radially from the tip; since the electron trajectories initially follow the lines of force, they also diverge, and a highly magnified emission map of the emitter surface can be obtained, for instance, by making the anode a fluorescent screen. This constitutes a field emission microscope, invented by E. W. Müller in 1936. Since work function and hence emission are affected by adsorbed layers, the field emission microscope is very useful for studying adsorption, particularly surface diffusion. Field emitters are widely used in "ordinary" and scanning electron microscopes as high-brightness quasi-point sources of electrons, since emission occurs as if it originated from the center of the emitter cap. *See* ELECTRON MICROSCOPE.

Field emission can also occur from electrode asperities into insulating liquids, and thus isolated atoms or molecules in high fields, and is then called field ionization. Field ionization forms the basis of the field ion microscope, and is a useful method of generating ions in analytical mass spectrometry. Internal field emission can occur from the valence to the conduction band of a semiconductor in a high field, and is then known as Zener breakdown. *See* ELECTRON EMISSION; FIELD-EMISSION MICROSCOPY. Robert Gomer

Field-emission microscopy

A technique that uses field emission of electrons or positive ions from a needle-shaped emitter to produce a magnified image of the emitter surface on a fluorescent screen. In the field electron microscope, the image reveals the variation in work function of the emitter surface. Due to the large lateral velocity of the emitted electrons, a resolution of only about 2.5 nanometers can be achieved. The large lateral velocity arises from a diffraction effect of the de Broglie wave and the large kinetic energy of electrons inside the metal; these effects are intrinsic quantum properties of particles and cannot be eliminated. The field electron microscope has been used to study adsorption and desorption of gases and vapor-deposited materials, surface migration of adsorption layers and absorbed atoms on single crystal faces, and surface reactions in catalysis. Medium-sized individual molecules such as phthalocyanin have been made visible also. *See* DE BROGLIE WAVELENGTH; ELECTRON MICROSCOPE; FIELD EMISSION; WORK FUNCTION (ELECTRONICS).

Field ion microscope. In the field ion microscope, the emitter is kept at a high positive potential while the microscope chamber is filled with helium, neon, or argon at a pressure of 10^{-5} to 10^{-4} torr (10^{-3} to 10^{-2} pascal) [**Fig. 1**]. Under a field of several tens of volts per nanometer, an image gas molecule can be ionized above a protruding surface atom when an atomic electron tunnels into the metal. The ion is then accelerated to the screen by the applied field. Every second about 1000 ions are formed above the same surface atom; thus the atom is continuously imaged. When the tip is cooled down to the temperature of liquid hydrogen or nitrogen, the thermal energy of the image gas molecules is greatly reduced. In combination with their very short de Broglie wavelength, a resolution of about 0.25 nm can be achieved. This is sufficient to resolve the atomic structure of most surfaces. In addition, surface atoms can be evaporated by the applied field if it is gradually increased. Therefore the atomic structure of lattice defects inside the bulk can be revealed also. By using a channel plate for image intensification and image gases of high or low ionization energy, a wide range of metals from beryllium to uranium, semiconductors, and some compounds such as high-temperature oxide superconductors can be imaged. The field ion microscope (Fig. 1) can also be used as a field electron microscope by simply reversing the polarity of the tip voltage and operating it in high vacuum. *See* TUNNELING IN SOLIDS.

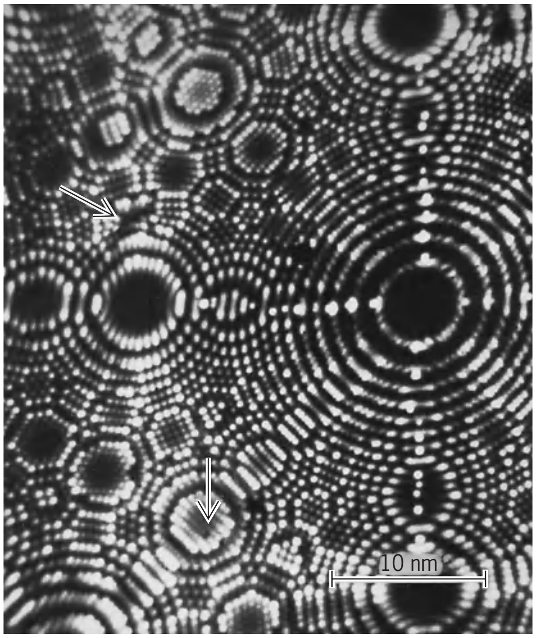


Fig. 2. Field ion micrograph of the atomic structure of an iridium crystal surface, with a vacancy and a dislocation shown by arrows, and many fully resolved net planes of atoms.

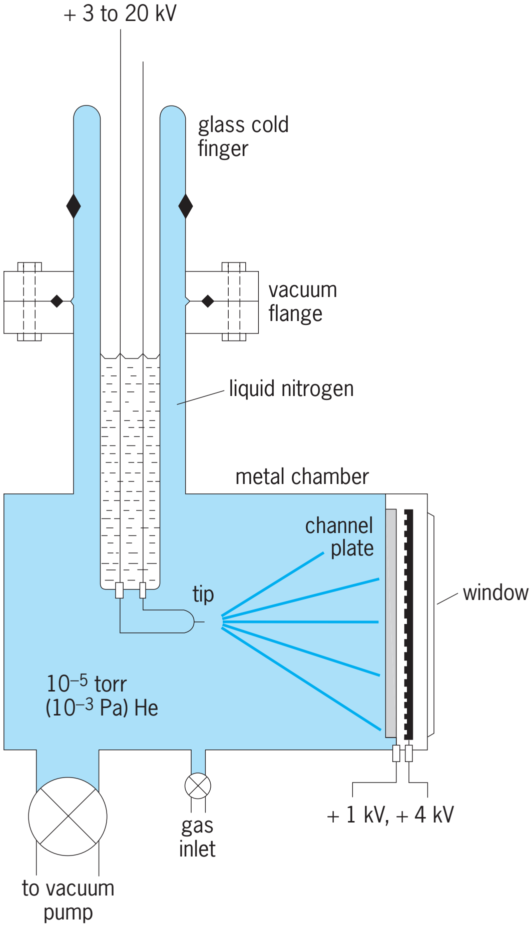


Fig. 1. Field ion microscope.

The field ion microscope has been used as a research tool for studying lattice defects such as vacancies (**Fig. 2**), interstitials, dislocations, grain boundaries, and radiation damages in metals and alloys. It has also been used to observe directly the behavior of single adsorbed atoms on metal surfaces. Random-walk diffusion of single atoms, interactions between single atoms, formation of atomic clusters from single atoms and their dissociation, and adsorption-layer superstructure formation have been observed and detailed atomic steps revealed. The field ion microscope has also been employed to study atomic reconstruction of solid surfaces and the formation of precipitates, oxide layers, thin silicide layers, and amorphous thin films. *See* CRYSTAL DEFECTS.

A slightly modified version of the field ion microscope is the field desorption microscope, which uses a double-channel plate. No image gas is needed; field desorbed ions form their own images. The desorption image reveals directly the spatial distribution of the adsorbed species on the surface. A shadowing technique has also been developed to image biological molecules. The field desorption images of adsorbed methane can reveal the shape of the biological molecules embedded beneath the methane layers.

Atom-probe field ion microscope. The instrument combines the field ion microscope with a single ion detection sensitivity mass spectrometer, usually a time-of-flight spectrometer. The tip is mounted on a gimbal system, and a channel plate screen assembly with a small probe hole is used. Behind the probe hole is a flight tube 3–24 ft (1–8 m) long (**Fig. 3**). The probe hole usually covers several atomic image diameters. The chemical identity of atoms chosen by the investigator from the field ion image can be analyzed one by one by adjusting the gimbal until the image of the atoms falls into the probe hole. Nanosecond high-voltage pulses or subnanosecond laser pulses are then applied to field-evaporate surface atoms one by one. Because of the ion optics, only those atoms that have their image covered by the probe hole can go through the probe hole and be detected. From the flight times of these ions, their mass-to-charge ratios are calculated and chemical species identified.

The atom-probe field ion microscope represents the ultimate in chemical analysis, that is, an atom selected by an investigator can be analyzed. As field evaporation starts from the edges of a surface layer, composition of a solid surface can also be analyzed atomic layer by atomic layer. The spatial resolution in atom-probe chemical analysis is approximately 5 to 10 angstroms in the lateral direction and much better than 1 angstrom in the vertical direction.

The three-dimensional atom probe uses a wide-angle position-sensitive ion detector for measuring the flight times of field-evaporated ions from a large area of the emitter surface. With this instrument, it is possible to analyze the three-dimensional elemental distribution of a material with a spatial resolution of approximately 0.5 nm.

The atom-probe field ion microscope has been successfully employed in surface science and materials science to study problems such as compositional variations in alloys due to segregation of alloy species and impurities to the surface, grain boundaries and phase boundaries, atomic exchange in surface diffusion, composition of small alloy phases such as precipitates and clusters, range profiles in particle irradiations, gas-surface interactions, and surface-catalyzed chemical reactions. The advantages of the atom-probe field ion microscope over other microanalytical instruments are the good spatial resolution and the high degree of sensitivity. Thus, very small precipitates and compositional modulations of subnanometer sizes in materials, where other instruments fail, can be analyzed with the atom-probe. *See* ALLOY STRUCTURES; GRAIN BOUNDARIES; SURFACE PHYSICS; TIME-OF-FLIGHT SPECTROMETERS.

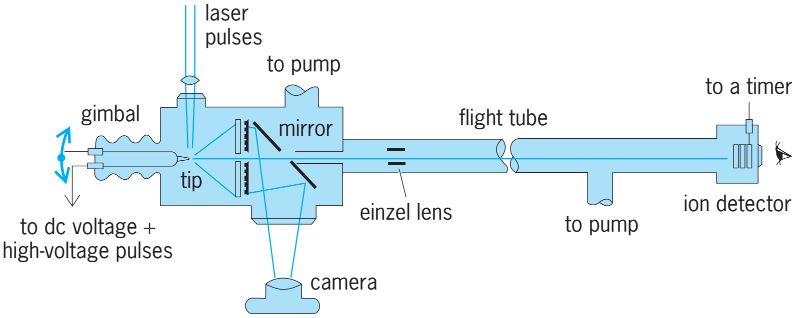


Fig. 3. Time-of-flight atom-probe field ion microscope. The einzel lens is an electrostatic one-piece lens for focusing the ion beam.

Tien Tzou Tsong

Bibliography. E. W. Müller and T. T. Tsong, *Field Ion Microscopy: Principles and Applications*, 1969; J. A. Panitz, Surface characterization by single-atom mass spectroscopy, *Crit. Rev. Solid State Sci.*, 5:153–178, 1975; T. T. Tsong, *Atom-Probe Field Ion Microscopy*, 1990; T. T. Tsong, Atom-probe field ion microscopy, *Phys. Today*, 46:24, 1993.

Field-programmable gate arrays

Very large scale integrated (VLSI) circuits that can be electrically programmed to implement complex digital logic. By combining many of the advantages of custom chips with the programmability of microprocessors, they are useful in a wide range of electronics applications. The name derives from their structure—a grid of programmable logic elements (gate array) and interconnection wires that can be configured at an engineers workbench, or even inside a consumer electronic device (programmed in the field), as opposed to other technologies that must be preconfigured during chip fabrication. *See* INTEGRATED CIRCUITS.

Design. Field-programmable gate arrays generally consist of a large number of programmable logic blocks embedded in a flexible routing structure (**Fig. 1**). The logic blocks implement the logic functions and flip-flops (simple memory elements that allow the system to remember past inputs and behaviors), while the routing structure wires these elements together into a complete system (**Figs. 2** and **3**). Typically, the logic blocks contain look-up tables, small 1-bit-wide memories that store the desired output for any possible combination of their inputs. To configure a look-up table to compute the function "return true if both A and B are true, or C is true," the table would have false for the three cases where C is false and at least one of A and B are false ($ABC = \{000, 010, 100\}$), and true for the other five cases. In a similar manner, any 3-input combinational Boolean function can be implemented in a 3-input look-up

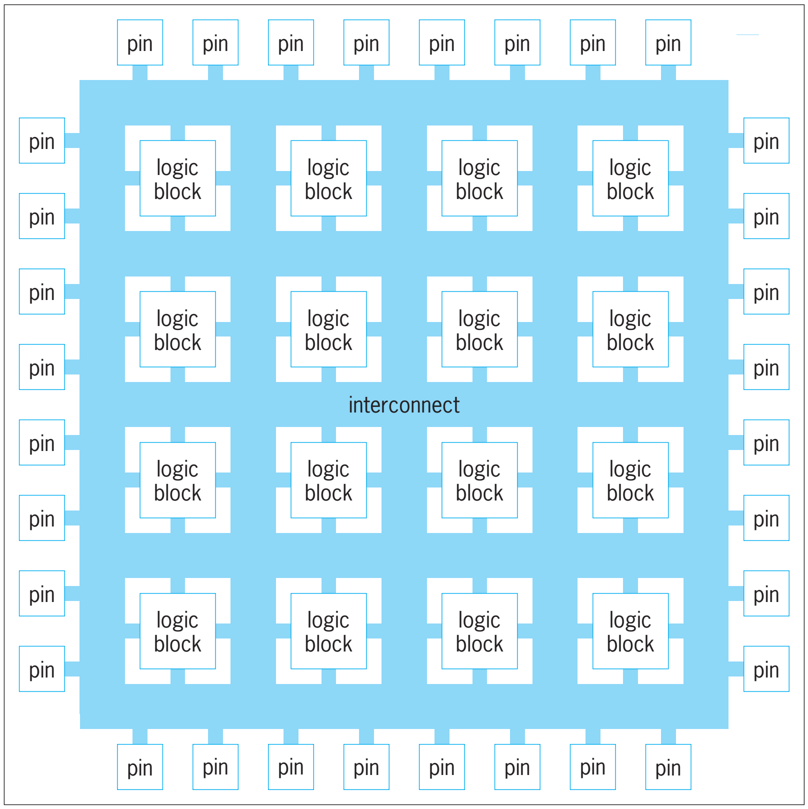


Fig. 1. Basic structure of a typical field-programmable gate array.

table. *See* LOGIC CIRCUITS; SEMICONDUCTOR MEMORIES.

Look-up tables are usually coupled with an optional D flip-flop (Fig. 2). All D flip-flops in the field-programmable gate array receive a periodic signal called a "clock," which transitions millions to hundreds of millions of cycles a second. Every time the clock changes from true to false, the D flip-flop's input D is captured and sent on to the output Q. In this way, the global clock signal synchronizes the operation of all memory elements in the system, and allows the design to remember and react to past inputs and behaviors.

The muxiplexor (Fig. 2) picks between the outputs of the look-up table and the flip-flop. If the programming bit P stores a 0 (false) value, the look-up table's output is sent to the output of the logic block; while if the programming bit is 1 (true), the D flip-flop's output is sent. In this way, the logic block can be programmed to compute a clocked or unclocked function.

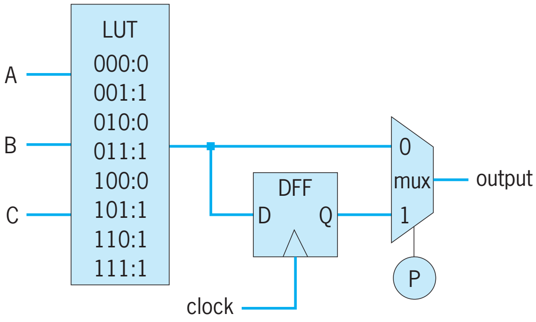


Fig. 2. Abstract representation of a logic block in a typical field-programmable gate array. LUT = Look-up table. DFF = D flip-flop (input D; output Q). Mux = muxiplexor (programming bit P).

The routing structure provides fairly flexible wiring connections between the outputs and inputs of the logic blocks on the chip. This allows the relatively simple logic blocks to be combined to support a complex computation, perhaps computing a function equivalent to a million or more basic 2-bit Boolean functions. The routing structure also connects to the chip input and output pins. The routing wiring pattern is generally fixed once the field-programmable gate array is programmed, instead of the dynamically changing communication pattern found in a typical computer network.

Modern field-programmable gate arrays often augment these basic elements with more complex units. Logic blocks have elements to speed up additions and subtractions of large values, while other parts of the array replace the logic blocks with more special-purpose elements. These might include multipliers, memories, and even complete microprocessors. Also, the input/output (I/O) pins often include the special hardware needed to support high-speed communication protocols to other chips or communication networks.

Programming. All of these elements within the field-programmable gate array are electrically programmable. The chip contains huge numbers of programming bits, individual elements that optionally turn on or off different features in the array. For example, each entry in a look-up table will be a bit programmable to represent either a true or a false. These programming bits might be implemented using antifuse, SRAM (static random-access memory), or EEPROM/Flash (electrically erasable programmable read-only memory/flash memory) technology. Antifuse programming points are normally disconnected wires in the array that can be "blown" by a high voltage to make a permanent electrical connection. SRAM are memory bits like the cells in a computer memory; they can hold a value for as long as power is maintained to the chip, but can also be changed at will. EEPROM/Flash is a hybrid of the two, based on isolated capacitors that can hold a value without power being applied, yet can be reprogrammed via high voltages.

The basic programming technology of a field-programmable gate array affects the way that it can be used. Chips based upon SRAM technology generally must have some external memory to hold the configurations, since they lose their values when the power is turned off; antifuse and EEPROM/Flash do not. Antifuse devices can be programmed only once; SRAM or EEPROM/Flash devices can be reconfigured to fix problems or change configurations. Because of this reconfigurability, SRAM-based and EEPROM/Flash-based FPGAs will be referred to as reconfigurable field-programmable gate array here. Overall, SRAM-based devices are generally the most successful commercially, since they do not need special fabrication techniques and thus can use the latest chip technologies.

To configure a field-programmable gate array for a specific circuit, a designer typically starts with code written in a hardware description language such as Verilog or VHDL. These are languages similar to C or Ada but are optimized for specifying hardware designs. This code is then converted to basic logic gates, such as ANDs and ORs, by a logic synthesis tool. A "technology-mapper" takes these logic gates and groups them into chunks that fit in the look-up tables in the field-programmable gate array. Finally, these look-up-table size groupings are assigned into specific logic blocks and routed together by the "placement" and "routing" tools respectively. At the end of this process, a file with the assignments to the field-programmable gate array's configuration bits (known as a bitstream) is generated. When this file is downloaded into a field-programmable gate array, it is configured to implement the user's design. Each of these steps is highly automated by software tools. *See* PROGRAMMING LANGUAGES.

Applications. Field-programmable gate arrays can be viewed as occupying a middle ground between programmable processors and custom chips. Compared to custom chips developed for a specific application (application-specific integrated circuits, or ASICs), field-programmable gate arrays are generally slower, are less power-efficient, and have a smaller logic capacity. However, an ASIC requires a much greater design time and a large up-front cost to set up fabrication for the first chip. Furthermore, unlike reconfigurable field-programmable gate arrays, ASICs cannot be changed after fabrication to handle bug fixes or functionality upgrades.

Compared to programmable processors, such as microprocessors and digital signal processors, field-programmable gate arrays can be faster, denser, and more power-efficient. However, field-programmable gate arrays are only more efficient than processors on tasks involving long repetitive tasks, such as multimedia processing and networking. Also, designing for field-programmable gate arrays is more complex than for microprocessors, and field-programmable gate arrays usually cannot support running multiple applications simultaneously (multitasking). *See* MICROPROCESSOR.

Although a microprocessor-based solution is usually preferred over an implementation with either an ASIC or a field-programmable gate array, some applications have speed or power requirements that make programmable processors unsuitable. Because of their relative strengths versus ASICs, field-programmable gate arrays are typically used in the following applications.

Low- to medium-volume hardware. Because ASICs have very high initial costs that must be amortized over each system sold, custom hardware may be affordable only for products that will sell a large number of units. And while in the past field-programmable gate arrays were able only to implement small parts of an application, their significant increase in capacity have allowed field-programmable gate arrays to now support complete systems in a single chip.

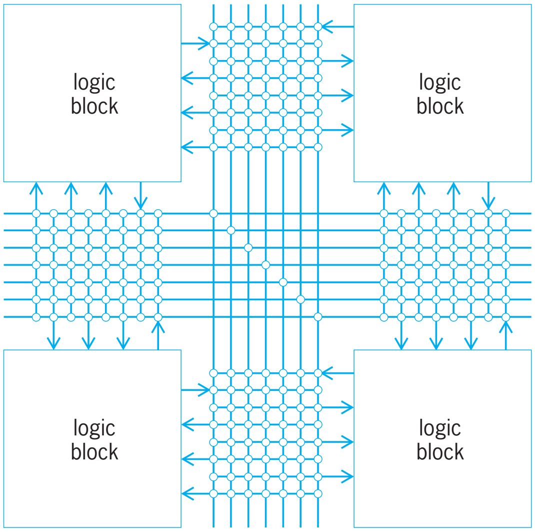


Fig. 3. Abstract representation of the interconnect structure for a typical field-programmable gate array. The circles are programmable connections between crossing wires.

Fast time-to-market. Even for applications with enough sales volume to justify the initial costs of an ASIC, the design time may be too long. Field-programmable gate arrays, which require less design time and have no extra delays for custom fabrication, can help get a product to market much faster than an ASIC.

Rapid prototyping. When a designer is developing a new circuit, it is often crucial to create a working prototype of the system for testing. This prototype will need to be changed multiple times as bugs and limitations are found. While ASICs cannot be changed, the reprogrammability of reconfigurable field-programmable gate arrays makes them ideal for prototyping.

Upgradability. For some applications, it is clear the circuitry will change after the system is sold. This includes devices supporting evolving protocols, complex designs susceptible to bugs, and applications where new features might be added over time. Reconfigurable field-programmable gate arrays can easily handle such systems, with the logic changed simply by providing a new bitstream.

Multimode systems. Some systems will need to support many different applications. For example, a single cellphone design might need to handle multiple protocols, depending on the country and the cellphone network provider. A field-programmable gate array can implement each protocol simply by providing it with the appropriate configuration file.

Run-time reconfiguration. A more aggressive use of a field-programmable gate arrays reconfigurability is to change the loaded configuration rapidly during the operation of the system. A field-programmable gate array otherwise is too small to implement a large computation but can do so if that computation is broken down into multiple, smaller pieces. This significantly increases the virtual capacity of the field-programmable gate array. ASICs would have to implement all of the configurations simultaneously.

Field-programmable gate arrays were introduced commercially in 1985, then became a multibillion-dollar industry. Due to their reprogrammability, ease of use, and low cost of hardware and support software, SRAM-based field-programmable gate arrays are one of the best hardware technologies for the novice or hobbyist. Scott A. Hauck

Bibliography. S. Brown and Z. Vranesic, *Fundamentals of Digital Logic with Verilog Design*, McGraw-Hill, 2002; K. Compton and S. Hauck, Reconfigurable computing: A survey of systems and software, *ACM Comput. Surveys*, 34(2):171–210, June 2002; S. Hauck, The roles of FPGAs in reprogrammable systems, *Proc. IEEE*, 86(4):615–639, April 1998; J. Villasenor and W. H. Mangione-Smith, Configurable computing, *Sci. Amer.*, pp. 66–71, June 1997.

Field theory (mathematics)

In algebra, the term field is used to designate an algebraic system or structure containing at least two elements and having two binary rules of composition: addition and multiplication (that is, if a and b are any two elements of the field, then $a + b$ and ab are defined and are elements of the field). The structure rules are as follows: The elements form an abelian (commutative) group under addition with the additive identity denoted by 0; that is, $a + 0 = a$ for all elements a. The set of nonzero elements (and there are some since the field has at least two elements) form an abelian group under multiplication with the multiplicative identity denoted by 1. It follows that all nonzero elements have a multiplicative inverse or reciprocal. The two rules of composition are related by the distributive law: $(a + b)c = ac + bc$ for all elements a, b, c. It follows from the distributive law that $a \cdot 0 = 0$ for all elements a, since $1 \cdot a = (1 + 0)a = 1 \cdot a + 0 \cdot a$, whence $0 = 0 \cdot a$. *See* GROUP THEORY.

The most familiar example of a field is the set **Q** of all rational numbers (0, 1, 2, $^1/_2$, −1, −2, $-^1/_2$, 3, $^1/_3$, $^2/_3$, . . .) with the usual rules for equality, addition, and multiplication. Other examples are the set **R** of all real numbers and the set **C** of all complex numbers (both with the usual rules of addition and multiplication). The set of rational functions (formal quotients of two polynomials with rational coefficients) also constitute a field under the usual rules for equality, addition, and multiplication. *See* ALGEBRA; COMPLEX NUMBERS AND COMPLEX VARIABLES; REAL VARIABLE.

A field satisfies all the axioms for a commutative ring with unity and in addition has the property that it contains the reciprocal of each nonzero element. A field is also an integral domain; in particular, if $ab = 0$ and $a \neq 0$, then $b = (a^{-1}a)b = a^{-1}(ab) = a^{-1} \cdot 0 = 0$. *See* RING THEORY.

There are fields that contain only finitely many elements. The simplest example is obtained by considering the residue class ring of the integers modulo a prime p. This field, which may be denoted by $\mathbf{F}_p$, has as its elements the p distinct arithmetic progressions or residue classes modulo the prime p, namely $I_0, I_1, \ldots, I_{p-1}$, where I_j consists of the integers j, $j \pm p$, $j \pm 2p$, ... The rules of composition are given by Eqs. (1) and (2), and can be used to construct addition and multiplication tables (**Table 1**).

$$I_j + I_k = I_s \text{ if } j + k \text{ lies in } I_s \tag{1}$$

$$I_j I_k = I_t \text{ if } jk \text{ lies in } I_t \tag{2}$$

The progression I_0 is the additive identity since $I_0 + I_k = I_k$ for all k, and I_1 is the multiplicative identity since $I_1 I_k = I_k$ for all k. In $\mathbf{F}_p$, if any I_k is added to itself p times, I_0 is obtained. The residue class rings modulo a composite number do not form a field, since in those rings some nonzero elements do not have multiplicative inverses.

Subfields and extensions. If F is a field and a subset A of F is a field under the rules of composition of F, then A is said to be a subfield of F and F is said to be an extension of A. Thus the field of rational numbers **Q** is a subfield of the field of real numbers **R** and **R** is an extension of **Q**. If a field F is an extension of the field A, then F is a vector space over A. If, as a vector space, F is of finite dimension n over A, then F is said to be an extension of A of degree n. In particular, **C** is an extension of **R** of degree 2. *See* LINEAR ALGEBRA.

Two fields F and F' are said to be isomorphic if there is a one-to-one mapping f of the elements of F onto the elements of F' such that $f(a + b) = f(a) + f(b)$ and $f(ab) = f(a)f(b)$ for all pairs of elements a, b in the field F. Every field F contains a unique subfield which is isomorphic either to the field **Q** of rational numbers or to $\mathbf{F}_p$ for some prime p. The finite field $\mathbf{F}_p$ occurs when repeated addition of 1 to itself gives 0. It follows that if F is a finite field, then it is an extension of $\mathbf{F}_p$, for some prime p, and hence is a vector space of finite dimension over $\mathbf{F}_p$. Thus the cardinality, or number of elements, of a finite field must always be a power of a prime. All

TABLE 1. Addition and multiplication tables for the residue class ring of the integers modulo p = 5

+	I_0	I_1	I_2	I_3	I_4	×	I_0	I_1	I_2	I_3	I_4
I_0	I_0	I_1	I_2	I_3	I_4	I_0	I_0	I_0	I_0	I_0	I_0
I_1	I_1	I_2	I_3	I_4	I_0	I_1	I_0	I_1	I_2	I_3	I_4
I_2	I_2	I_3	I_4	I_0	I_1	I_2	I_0	I_2	I_4	I_1	I_3
I_3	I_3	I_4	I_0	I_1	I_2	I_3	I_0	I_3	I_1	I_4	I_2
I_4	I_4	I_0	I_1	I_2	I_3	I_4	I_0	I_4	I_3	I_2	I_1

TABLE 2. Addition and multiplication tables for the field *J* of four elements

$+$	0	1	a	$a+1$
0	0	1	a	$a+1$
1	1	0	$a+1$	a
a	a	$a+1$	0	1
$a+1$	$a+1$	a	1	0

$\times$	0	1	a	$a+1$
0	0	0	0	0
1	0	1	a	$a+1$
a	0	a	$a+1$	1
$a+1$	0	$a+1$	1	a

finite fields of the same cardinality are isomorphic to each other. For example, a field of four elements is isomorphic to the field J (**Table 2**), which contains $\mathbf{F}_2$ and is an extension of $\mathbf{F}_2$. In a finite field there exists an element whose powers give all nonzero elements. In the case of $\mathbf{F}_5$, both I_2 and I_3 are such elements. In the field J (Table 2), there are two such elements, which may be designated a and $a+1$.

Given an integral domain D, there exists (to within isomorphism) a unique smallest field $Q(D)$ containing D. The field $Q(D)$ is called the quotient field of D, and it stands in relation to D as the field $\mathbf{Q}$ of rational numbers stands in relation to the ring of integers; indeed $\mathbf{Q}$ is the quotient field of the ring of integers.

Solutions of polynomial equations. Every linear equation $ax+b=0$ with $a \neq 0$, and a, b in a field F has a solution in F, namely $-(b/a)$. But this need not be the case for equations of higher degree; for example; neither $x^2-2=0$ nor $x^2+1=0$ has a solution in the field $\mathbf{Q}$ of rational numbers. Given a field F and a polynomial $p(x) = a_0x^n + \cdots + a_n$, $n > 0$, with $a_0 \neq 0, a_1, \ldots, a_n$ in F, there exist extensions of F containing a solution for $p(x) = 0$. For example, the set B, consisting of all elements of the form $a + b\sqrt{2}$, where a and b are rational numbers, with the composition rules $(a+b\sqrt{2}) + (c+d\sqrt{2}) = (a+c) + (b+d)\sqrt{2}$ and $(a+b\sqrt{2})(c+d\sqrt{2}) = (ac + 2bd) + (ad+bc)\sqrt{2}$ is a field which extends the field of rational numbers and contains a solution of the polynomial equation $x^2-2=0$. The field J contains a root of the polynomial x^2+x+1, which is inducible over $\mathbf{F}_2$. If $p(x)$ is a polynomial with coefficients in a field F and if $p(x)$ is irreducible over F, that is, $p(x)$ does not factor into a product of polynomials with coefficients in F of lower degree, then there is a unique (to within isomorphism) extension K of F containing a solution for the polynomial equation $p(x) = 0$. If the degree of a field A over F is n, then every element of A satisfies an irreducible polynomial over F of degree at most n.

If a field F is such that every polynomial equation over F of positive degree has a solution in F, then F is said to be algebraically closed. The field of all complex numbers is algebraically closed; but this is a theorem in analysis. The field of rational numbers, the fields of residue classes modulo a prime p, and the field of real numbers are not algebraically closed. Given any field F, there exists a unique (to within isomorphism) minimal extension $\tilde{F}$ of F such that $\tilde{F}$ is algebraically closed. If the field F is countable (that is, can be put into a one-to-one correspondence with the positive integers), then $\tilde{F}$ is countable. Since the field $\mathbf{Q}$ of rational numbers and the fields $\mathbf{F}_p$ are countable, it follows that $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{F}}_p$ are countable. In particular, it follows that $\tilde{\mathbf{Q}} \neq \mathbf{C}$, since $\mathbf{C}$ is not countable. It is easily seen that $\mathbf{C}$, the field of complex numbers, is isomorphic to $\tilde{\mathbf{R}}$, the algebraic closure of the field of real numbers.

If K is an extension of a field F, an element of K is said to be algebraic over F if that element is a solution of a polynomial equation with coefficients in the field F. If an element of the larger field K is not algebraic over F, the element is said to be transcendental over F. The real numbers $\sqrt{2}$ and $\sqrt[3]{1+3\sqrt{5}}$ are algebraic over the field of rational numbers, while the numbers e, π, and

$$\theta = \sum_{n=1}^{\infty} 10^{-n!}$$

are transcendental over the field of rational numbers. The proofs are not obvious and use analysis. It is an important and difficult problem in number theory to determine if a given real number is algebraic or transcendental over the field of rational numbers. If an element η of K is transcendental over the field F, then the smallest field containing F and η is isomorphic to the field of rational functions in one variable over F; that is, the quotient field of the ring of polynomials $F[t]$ with coefficients in F. If all the elements of some extension K containing the field $\mathbf{Q}$ of rational numbers has all its elements algebraic over $\mathbf{Q}$, then one seeks to study the arithmetic structure of K; for example, to answer whether K has elements which uniquely generate its multiplicative group as the primes generate the multiplicative group of nonzero rationals. *See* CIRCLE; E (MATHEMATICS); NUMBER THEORY; ROOT (MATHEMATICS).

Valuations. An important class of fields are those which possess a valuation. The simplest valuations are functions v from the field to the nonnegative real numbers with the properties: [i] $v(a) = 0$ if and only if $a = 0$, [ii] $v(a+b) \leq v(a) + v(b)$, [iii] $v(ab) = v(a)v(b)$. The ordinary absolute value is a valuation on the fields $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{C}$. The field $\mathbf{Q}$ has other valuations. For example, if p is a prime, write $a/b = p^{\alpha}(c/d)$, where α is an integer and p does not divide cd; then $v_p(a/b) = p^{-\alpha}$ is a valuation on $\mathbf{Q}$ known as the p-adic valuation. When a field has a valuation, one can study its completion in the sense of a corresponding topology. The fields $\mathbf{R}$ and $\mathbf{C}$ are complete under the ordinary absolute value and $\mathbf{R}$ is the completion of $\mathbf{Q}$ under this valuation. The completion of $\mathbf{Q}$ under v_p creates the fields of p-adic numbers, $\mathbf{Q}_p$. The fields $\mathbf{Q}_p$ and their extensions occur in the study of arithmetic properties of fields algebraic over $\mathbf{Q}$. *See* TOPOLOGY.

Automorphisms of fields. A very useful tool in the study of algebraic structures is the group of automorphisms of the structure. An automorphism of a field is an isomorphism of a field onto itself. The study of relations between extensions of a field by means of automorphism groups is known as Galois theory. Important problems can often be formulated as questions about solutions of the polynomial equation, and these in turn can be resolved by considering the automorphism group of the minimal extension containing all solutions of the equation. Among such problems are the question of squaring the circle and ruler and compass constructions. D. J. Lewis

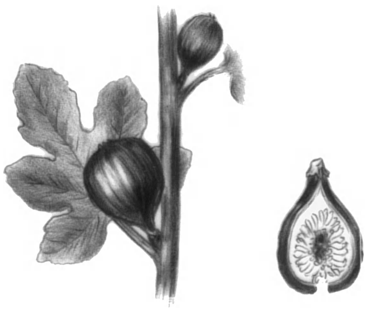
Bibliography. M. Artin, *Algebra*, 1995; L. N. Childs, *A Concrete Introduction to Higher Algebra*, 2d ed., 2000; J. B. Fraleigh, *A First Course in Abstract Algebra*, 6th ed., 1998; I. N. Herstein, *Abstract Algebra*, 3d ed., 1996; N. Jacobson, *Lectures in Abstract Algebra*, vol. 3, 1964, reprint 1981; S. Lang, *Algebra*, 3d ed., 1993; O. Zariski and P. Samuel, *Commutative Algebra*, 1975, reprint 1986, 1991.

Fig

A species of deciduous tree, *Ficus carica*, of the mulberry family (Moraceae). Leaves are characteristically palmate, with 3 to 7 lobes. A milky latex containing the protein-digesting enyzme ficin is produced. It is of southwest Asian origin and is now cultivated in the subtropical regions of the world for its edible fruit, especially in Asia, Africa, Europe, and the United States.

Flowers and fruit. The fig is a multiple or collective fruit (see **illus.**). It develops from a fleshy compound inflorescence in the form of a hollow structure (syconium) with an apical canal or pore (ostiole) and numerous tiny unisexual flowers lining the inner wall. The flowers are long-style pistillate (female); short-style pistillate (female); and staminate (male). They are significant in the life history of a fig wasp species, *Blastophaga psenes*. The true fruits are tiny, juicy stone fruits (drupelets) that develop from the ovaries of the pistillate flowers. *See* FLOWER; FRUIT.

Tree forms. Two forms of trees characterize *F. carica*: fig and caprifig. The fig bears long-style pistillate flowers exclusively and produces the edible figs of commerce. Depending on cultivar, it may produce two crops annually: a spring crop, commonly called the first or breba crop, and a summer crop, commonly called the main crop or the second crop, even if there is no first crop.



***Ficus carica*, details of branch and fruit.**

The caprifig bears short-style pistillate flowers and staminate flowers which are clustered around the inner opening of the ostiole. It produces three crops annually: an overwinter crop, which is initiated in fall and matures in early spring (the mamme crop); a crop that is initiated in early spring and matures in late spring or early summer (the profichi crop); and a crop that is initiated in summer and matures in fall (the mammoni crop). Caprifigs are the source of pollen for certain cultivars of figs whose flowers must be pollinated to set a crop of fruit.

Cultivars. More than 600 cultivars of figs have been named. Depending on the kinds of crops they bear, they are classified into three horticultural types: common, San Pedro, and Smyrna. Trees of the common type produce good summer crops. Depending on cultivar and climatic conditions, they may or may not produce a spring crop. Caprification is not necessary for setting a crop. Important cultivars, along with names by which they are better known in the trade, are Brunswick (Magnolia), Dottato (Kadota), Franciscana (Mission), Genoa, Malta (Celeste), San Piero (Brown Turkey), and Verdone (Adriatic).

Trees of the San Pedro type generally produce good spring crops and poor summer crops, if any. The latter may or may not need to be caprified to set a crop. Important cultivars of this type are San Pedro, Blanquette, Dauphine, King, and Lamperia. Trees of the Smyrna type seldom bear a spring crop. They set good summer crops, but only if the syconia are caprified, that is, if the pistillate flowers are pollinated with subsequent development of the ovaries into drupelets. Important cultivars are Sari Lop (Lob Injir or Calimyrna) and Kalamata. Fig cultivars are propagated almost entirely by the rooting of cuttings.

Caprification. Caprification is effected by the females of the fig wasp, which colonizes caprifig trees. A wasp develops in the ovule of each pistillate flower in which an egg has been laid. The females become covered with pollen when they crawl through the staminate flowers before leaving the syconium through the ostiole. They enter the fig syconia through the ostiole and pollinate the flowers as they crawl over them while attempting to lay eggs in the ovules. However, they fail in the attempts to oviposit because the styles of the flowers are longer than the ovipositor. This procedure takes place only in late spring and early summer, when the profichi crop of the caprifig trees ripens and the flowers of the fig trees are receptive to pollination.

Harvesting and utilization. Figs that are to be eaten or marketed fresh, canned, candied, pickled, or made into jam or preserve usually are picked from the trees by hand. The cultivars most widely used are Brunswick, Franciscana, King, Malta, San Piero, and Verdone.

The bulk of the world's fig crop either is eaten dry or is marketed as dried figs or as fig paste, which is used for making fig bars and other bakery products. Figs to be used in these ways usually are allowed to dry on the trees and drop to the ground for harvesting either by hand or by machinery. The cultivars grown most widely for dried figs are the following: in California, Sari Lop, Dottato, Franciscana, and Verdone; in Greece, Kalamata; in Italy, Dottato; in Turkey, Sari Lop. *See* FRUIT, TREE. William B. Storey

Bibliography. I. J. Condit, *The Fig*, 1947; A. J. Downing, *Fruit and Fruit Trees of America*, 1993; J. Galil, Fig biology, *Endeavour*, 1(2):52–56, 1977.

Filbert

A nut from any plant in the genus *Corylus*; also called hazelnut. About 15 species are recognized, distributed widely over the North Temperate Zone and ranging in size from shrubs to tall trees. Filberts in commerce are derived mostly from the European species *C. avellana* and *C. maxima*. Hybrids of these with the American species, *C. americana*, are grown sparingly in the eastern United States.

Filberts belong in the birch family (Betulaceae), which produces the staminate (male) flowers in catkins on the previous season's growth and the pistillate (female) flowers, which form the nuts, at the end of the current season's shoots. The windblown pollen is shed from the catkins in late winter or early spring, when the pistillate flowers are receptive to fertilization. Cross pollination is necessary to produce a crop of nuts. The round or oval nuts, $^1/_2$ to $^3/_4$ in. (1.3 to 1.9 cm) in diameter, are developed in a leaflike husk, from which they fall at maturity.

Filberts are grown widely in Europe and the Mediterranean basin, where much of the crop is consumed locally. Imports to the United States come mostly from Turkey. In the United States filberts are grown in Oregon and Washington.

American filberts are grown as small trees with single trunks. This requires frequent removal of the sucker shoots, which normally grow from the base and produce a bush. In Oregon harvesting is done with machines which rake the nuts into windrows and pick them up from the smooth-rolled ground.

Filberts are sold in shell, or the kernels are roasted and used in confectionery and baked goods. *See* NUT CROP CULTURE. Laurence H. MacDaniels

Film (chemistry)

A material in which one spatial dimension, thickness, is much smaller than the other two. Films can be conveniently classified as those that support themselves and those that exist only as layers on top of a supporting substrate. The latter are known as thin films and have their own specialized science and technology.

Thin films. Thin films, from one to several hundred molecular layers, are generally defined as those that lie on a substrate, either liquid or solid. Thin films may form involuntarily at a surface or they may be created by some external agency.

Involuntary formation. The molecules at the surface of a liquid or solid are subjected to an imbalance of intermolecular attractions. At a liquid surface, this imbalance gives rise to the phenomenon of surface tension and all its related effects known as capillarity. At a solid surface, the lattice is distorted to reduce its free energy, to a degree determined by the polarizability of the atoms or ions and the nature of the contiguous phase. Compounds such as mercury(II) iodide (HgI_2), lead iodide (PbI_2), and molybdenum sulfide (MoS_2), or elemental sulfur develop such low surface energies by this process that they become naturally hydrophobic. Even water-soluble salts such as chromium chloride ($CrCl_3$) may develop a hydrophobic surface. Other substances, composed of less polarizable atoms, continue to reduce their surface energies spontaneously by adsorbing molecules from the atmosphere; this is an example of a thin film formed involuntarily. Another example of the rearrangement of a surface layer is the presence of a thermodynamically stable liquid film on the surface of ice, even at temperatures 5–10°C (9–18°F) below the melting point, forming a transitional zone between the bulk crystalline phase of the ice and the surrounding air. Clearly, this layer, which must be a few hundred molecular diameters in thickness, cannot be the same as ordinary water. *See* ADSORPTION; INTERFACE OF PHASES; SURFACE TENSION.

Solutes that lower the interfacial free energy are said to be surface-active and are known as surfactants. Molecules of such solutes are more concentrated in the surface layer than in the bulk phase. Well-known examples in water are the soaps and detergents. These are examples of spontaneously formed films. The phenomenon is more familiar in liquid solutions but it is also known to occur in solid solutions. *See* DETERGENT; SOAP; SOLID SOLUTION; SURFACTANT.

Formation by external action. Monomolecular films on the surface of water can be made by adding appropriate materials in extremely small quantity to the surface. These materials are inappreciably soluble in water, rendered so by a large hydrocarbon or fluorocarbon functional group on the molecule or by other means; they are, however, able to dissolve in the surface of water by virtue of a hydrogen-bonding group or groups on the molecule. These monolayers can be lifted off the surface of water and deposited onto a solid substrate, such as glass or platinum. By repeated deposition on top of underlying layers, builtup films that are many monolayers thick can be obtained. These are known as Langmuir-Blodgett films; they occur in biological membranes and have applications in microlithography, integrated optics, electrooptics, electronic displays, photovoltaic cells, two-dimensional magnetic arrays, and field-effect devices. *See* MICROLITHOGRAPHY; MONOMOLECULAR FILM.

Films can also be deposited directly onto solid substrates by evaporation of material in a vacuum, by sputtering, by chemical reaction (chemical vapor deposition), by ion plating, or by electroplating. Suitable materials for vacuum deposition are metal alloys or binary inorganic compounds such as oxides, halides, sulfides, selenides, and tellurides. Some polymers are also suitable, but compounds that decompose on heating are not. This process is used extensively for thin-film metallization of glass and ceramics and in microelectronic circuitry. A modification of the method is atomic layer epitaxy, in which a chemisorbed monolayer is left after less strongly held molecules have been thermally desorbed. The process is repeated to build up successive chemisorbed layers, ultimately producing a monocrystalline deposit. This technique has been used to create a monocrystalline film of diamond on a silicon substrate, a development that may lead to a powerful new generation of computer chips. In sputtering, the material is removed from a source (the cathode), is carried by a plasma, and is deposited on a substrate (the anode). In ion plating, the material is evaporated in the positive glow region of a gas discharge at a pressure of a few pascals. A negative bias of several kilovolts with respect to the plasma is applied to the substrate. *See* CHEMICAL VAPOR DEPOSITION; DIAMOND; ELECTROPLATING OF METALS; PHYSICAL VAPOR DEPOSITION; SPUTTERING; VACUUM METALLURGY.

Thin films play an extensive role in both traditional and emerging technologies. Examples are foams and emulsions, the active layers in semiconductors, the luminescent and protective layers in electroluminescent thin-film displays, and imaging and photoelectric devices. *See* ELECTROLUMINESCENCE; ELECTRONIC DISPLAY; EMULSION; FOAM; SOLAR CELL.

Self-supported films. Films in this category have a nominal thickness not larger than 250 micrometers. Films of greater thickness are classified as sheets or foils.

Films are commonly composed of organic polymers, either thermoplastic resins or cellulose-based materials. The three produced in largest quantity are polyethylene, poly(vinyl chloride), and polystyrene. These are used chiefly in packaging applications. The chief requirements of such films are cost economy, transparency, uniform thickness, tear resistance, absence of pinholes, and inertness to moisture and chemicals. *See* POLYMER.

The lower limit of thickness of a self-supported film is about 25 μm. Gold leaf, which historically was the first such film to be produced, can be brought to a thickness of 0.01 μm, but at this degree of attenuation it is too fragile to exist without a supporting substrate. Sydney Ross

Bibliography. M. Ohring, *The Materials Science of Thin Films*, 2d ed., 2001; K. R. Osborn and W. A. Jenkins, *Plastic Films: Technology and Packaging Applications*, 1992; J. A. Venables, *Introduction to Surface and Thin Film Processes*, 2000.

Filtration

The separation of solid particles from a fluid-solids suspension of which they are a part by passage of most of the fluid through a septum or membrane that retains most of the solids on or within itself. The septum is called a filter medium, and the equipment assembly that holds the medium and provides space for the accumulated solids is called a filter. The fluid may be a gas or a liquid. The solid particles may be coarse or very fine, and their concentration in the suspension may be extremely low (a few parts per million) or quite high (>50%).

The object of filtration may be to purify the fluid by clarification or to recover clean, fluid-free particles, or both. In most filtrations, the solids-fluid separation is not perfect. In general, the closer the approach to perfection, the more costly the filtration; thus the operator of the process cannot justify a more thorough separation than is required.

Gas filtration. It is frequently necessary to remove solids (called dust) from a gas-solids mixture because (1) the dust is a contaminant rendering the gas unsafe or unfit for its intended use; (2) the dust particles will ultimately separate themselves from the suspension and create a nuisance; or (3) the solids are themselves a valuable product that in the course of its manufacture has been mixed with the gas. One method of separating solids from a gas is by gas filtration. Among other methods are centrifugal separation, infringement, scrubbing, and electrostatic precipitation.

Three kinds of gas filters are in common use. Granular-bed separators consist of beds of sand, carbon, or other particles which will trap the solids in a gas suspension that is passed through the bed. The bed depth can range from less than an inch to several feet. Bag filters are bags of woven fabric, felt, or paper through which the gas is forced; the solids are deposited on the wall of the bag. A simple example is the dust collector of a household vacuum cleaner. Air filters are light webs of fibers, often coated with a viscous liquid, through which air containing a low concentration of dust, normally less than 5 grains/1000 ft^3 (approximately 28 m^3), can be passed to cause entrapment of the dust particles. The filter installed in the air-intake lines of most household furnaces is an air filter. Gas filters can be used to remove particles in the size range from 1 nanometer to 100,000 nanometers. *See* DUST AND MIST COLLECTION.

Liquid filtration. Liquid-solids separations are important in the manufacture of chemicals, polymer products, medicines, beverages, and foods; in mineral processing; in water purification; in sewage disposal; in the chemistry laboratory; and in the operation of machines such as internal combustion engines. Filtration is the most versatile and widely used operation for such applications. In some cases, however, gravity or centrifugal sedimentation is preferred to filtration. *See* CENTRIFUGATION; MECHANICAL SEPARATION TECHNIQUES; THICKENING.

Liquid filters are of two major classes: cake filters and clarifying filters. The former are so called because they separate slurries carrying relatively large amounts of solids. They build up on the filter medium as a visible, removable cake which normally is discharged "dry" (that is, as a moist mass), frequently after being washed in the filter. It is on the surface of this cake that filtration takes place after the first layer is formed on the medium. The feed to cake filters normally contains at least 1% solids. Clarifying filters, on the other hand, normally receive suspensions containing less than 0.1% solids, which they remove by entrapment on or within the filter medium without any visible formation of cake. The solids are normally discharged by backwash or by being discarded with the medium when it is replaced.

Cake filters. There are three classes of cake filters, depending on the driving force producing the separation. In pressure filters, superatmospheric pressure is applied by pump or compressed gas to the feed slurry, and filtrate is discharged at atmospheric pressure, or higher. In vacuum filters, the pressure of the atmosphere is the driving force, a vacuum being applied to the filtrate collection chamber. In centrifugal filters, centrifugal force causes the filtrate to flow through the cake and filter medium.

Cake filters are also classified as batch-operating or continuous. In general, pressure filters are batch or intermittent filters, vacuum filters are continuous machines, and centrifugal filters may be either.

Pressure filters. There are three principal industrial classes of pressure filters: filter presses, horizontal plate filters, and leaf filters. They may be operated at 250 lb/in.^{2} (1724 kilopascals) or higher feed pressure.

Filter presses are the simplest and most versatile pressure filters. One form, the plate-and-frame press, is shown in **Fig. 1**. This filter consists of a number of rectangular or circular vertical plates, alternating with empty frames, assembled on horizontal rails. One fixed head and one movable head hold the plates and frames together. Both sides of each plate and the inner surfaces of the heads constitute the filtration area. The medium consists of filter cloths hung over the plates and covering all filter areas. The plates and frames are compressed by a powerful screw, so that the cloths also act as gaskets to prevent the filtrate from leaking between the plates. Filtrate collects on the plate surfaces and leaves the press through drainage channels; cake collects in the frames and is dumped out after the filtration cycle is finished.

The horizontal plate filter consists of a number of horizontal circular perforated drainage plates operating in parallel and placed one above another within a cylindrical shell (**Fig. 2**). A filter cloth or paper is placed on each plate, which has a drainage chamber beneath. During filtration, the whole shell is filled with slurry under filtering pressure. At the end of the cycle, the shell is drained and filter cake removed.

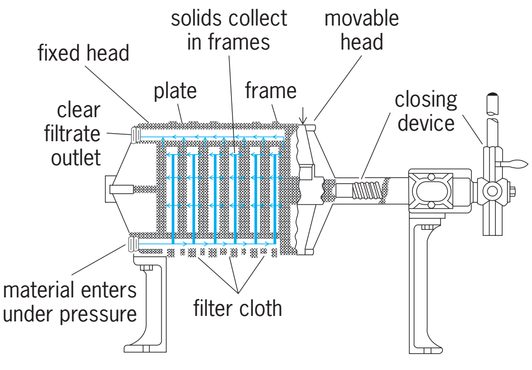


Fig. 1. Corner-feed, closed-discharge filter press, a type of plate-and-frame press. (*T. Shriver and Co.*)

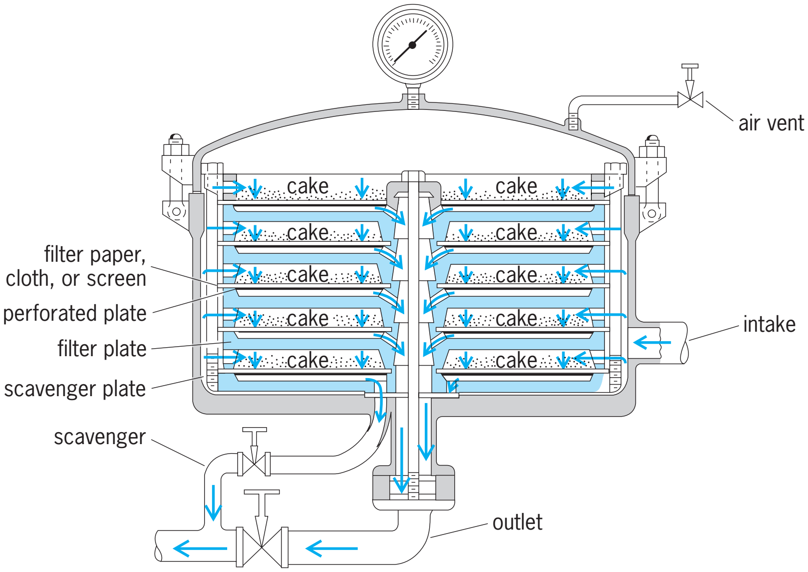


Fig. 2. Sparkler horizontal plate filter. (*After R. H. Perry and D. Green, eds., Perry's Chemical Engineers' Handbook, 6th ed., McGraw-Hill, 1984*)

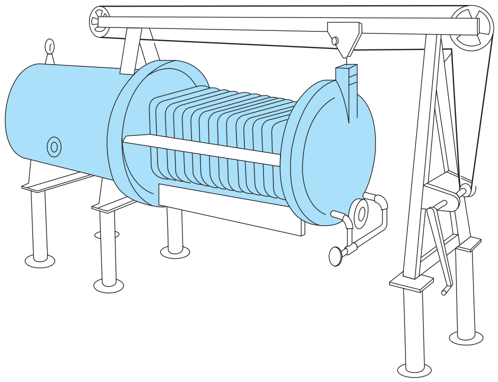


Fig. 3. Horizontal tank pressure leaf filter. (*After R. H. Perry and D. Green, eds., Perry's Chemical Engineers' Handbook, 6th ed., McGraw-Hill, 1984*)

Pressure leaf filters consist of pressure tanks filled with slurry into which a number of filter leaves connected to drainage pipes are immersed. The leaves may be grids covered with fitted cloth bags, shallow boxes of screen cloth, or porous cylindrical tubes of metal or porcelain. They may be horizontal or vertical. A horizontal tank filter with vertical leaves is shown open for cleaning in **Fig. 3**. When the filter cake can be flushed away instead of recovered dry, a pressure leaf filter can be operated for many cycles over a period of months without being opened. Such filters are adaptable to automatic programmed operation.

Continuous-vacuum filters. These filters are used principally for free-filtering solids. Their most important forms are rotary-drum, rotary-disk, horizontal belt, and horizontal pan or table filters.

The horizontal drum filter is a cloth-covered horizontal cylinder partially immersed in the slurry (**Fig. 4**). The periphery of the drum consists of a number of individual filter elements, each connected to one of the trunnions (drum supports) by pipes. The trunnion contains a special valve which applies suction to each element from the time it enters the slurry until just before it reaches the doctor blade. It then directs compressed air to the element until the element again submerges. Suction is then reestablished. Filtrate flows to the trunnion under the action of the vacuum and then out of the filter, as does wash water sprayed on the cake. The cake is sucked dry of wash and doctored off the filter.

In a disk filter, the rotating drum is replaced by an array of parallel disks, the faces of which are the cloth-covered filtering surface (**Fig. 5**). Each disk is divided into trapezoidal filter elements controlled by a valve, as in the drum filter.

Horizontal filters are useful when heavy solids are to be filtered or when thorough washing is required. They consist either of an endless belt or a rotating annular table, the top surface of each being the filtering surface. The belt passes over a suction box; the table is divided into filter elements connected to a valve, as in the drum filter. Instead of a table being used, the filter elements can be separated into individual trays, annularly arrayed, that can be tipped to assist in the cake dumping.

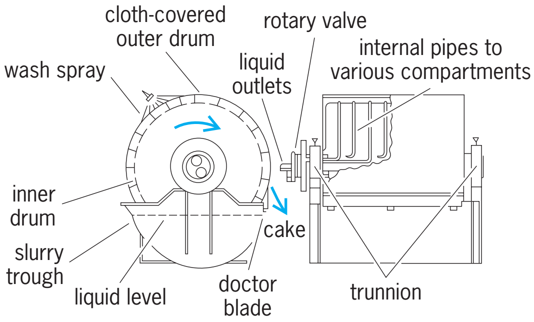


Fig. 4. Continuous vacuum rotary filter. (*After W. L. McCabe and J. C. Smith*, *Unit Operations of Chemical Engineering*, *McGraw-Hill*, *3d ed.*, *1976*)

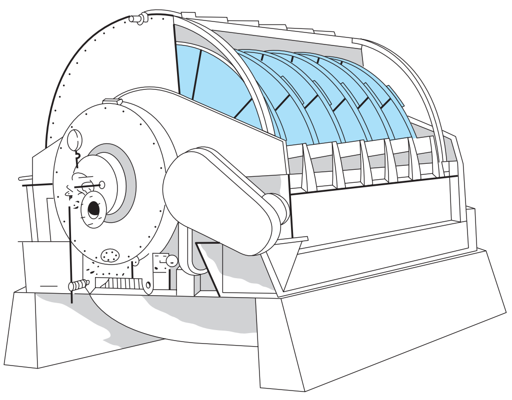


Fig. 5. American disk filter. (*After R. H. Perry and D. Green*, *eds.*, *Perry's Chemical Engineers' Handbook*, *6th ed.*, *McGraw-Hill*, *1984*)

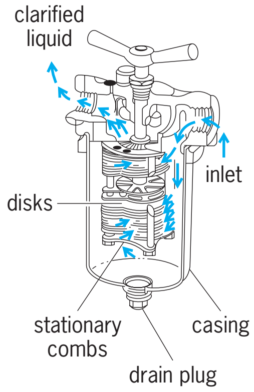


Fig. 6. Cartridge filter. (*After W. L. McCabe and J. C. Smith*, *Unit Operations of Chemical Engineering*, *McGraw-Hill*, *3d ed.*, *1976*)

Clarifying filters. When the object of filtration is to produce a solids-free liquid from a dilute suspension, a clarifying filter may be used. Alternatively, sedimentary clarifiers or centrifuges may be chosen. *See* CLARIFICATION; SEDIMENTATION (INDUSTRY).

Clarifying filters are pressure filters that employ a very fine filter medium, the pores of which are sufficiently small to prevent passage of particles of the size that must be removed. In cake filtration, cake particles bridge relatively large openings in the medium and create a new filtering surface; in clarifying filtration, each particle is collected on or within the filter medium.

Clarifying filters may be in the form of plate filters (analogous to shallow-frame filter presses), disk filters (analogous to horizontal plate filters), cartridges (analogous to leaf filters), and granular beds. The last-mentioned type is a special case; the principal example of the use of granular beds is in water filtration, commonly carried out by sand or coal filters operated by gravity instead of under pressure. A cartridge filter is shown in **Fig. 6**.

Although the filter medium is important in every filtration, it assumes paramount importance in clarifications. Aside from granular beds, there are four general types of clarifying mediums: spaced wires that strain out solids (lower limit, 25 micrometers); depth mediums, masses of fibers or interconnected pores that trap particles by interception (lower limit, 5 μm); micronic sheets or pads of carefully sized fibers (lower limit, 0.01 μm); and hyperfiltration membranes (lower limit, 1 nm). A special case of granular-bed clarification occurs when a cake of diatomaceous earth is laid down as a precoat on a coarse filter, thereby converting it to a clarifying filter. S. A. Miller

Bibliography. N. P. Cheremisinoff and D. S. Azbel, *Liquid Filtration for Process and Pollution Control*, 1989; W. L. McCabe, J. C. Smith, and P. Harriott, *Unit Operations of Chemical Engineering*, 7th ed., 2004; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; A. Rushton, A. S. Ward, and R. G. Holdich,

Solid-Liquid Filtration and Separation Technology, 2d ed., 2000; R. L. Wakeman (ed.), *Progress in Filtration and Separation*, vols. 1–4, 1986–1991.

Fine structure (spectral lines)

A term referring to the closely spaced groups of lines observed in the spectra of the lightest elements, notably hydrogen and helium. The components of any one such group are characterized by identical values of the principal quantum number n, but different values of the azimuthal quantum number l and the angular momentum quantum number j.

According to P. A. M. Dirac's relativistic quantum mechanics, those energy levels of a one-electron atom that have the same n and j coincide exactly, but are displaced from the values predicted by the simple Bohr theory by an amount proportional to the square of the fine-structure constant α. The constant α is dimensionless, and nearly equal to $^1/_{137}$. Its value is actually 0.007297352533±0.000000000027. In 1947 deviations from Dirac's theory were found, indicating that the level having $l = 0$ does not coincide with that having $l = 1$, but is shifted appreciably upward. This is the celebrated Lamb shift named for its discoverer, Willis Lamb, Jr. Modern quantum electrodynamics accounts for this shift as being due to the interaction of the electron with the zero-point fluctuations of the electromagnetic field. *See* ATOMIC STRUCTURE AND SPECTRA; QUANTUM ELECTRODYNAMICS; QUANTUM NUMBERS; RELATIVISTIC QUANTUM THEORY.

In atoms having several electrons, this fine structure becomes the multiplet structure resulting from spin-orbit coupling. This gives splittings of the terms and the spectral lines that are "fine" for the lightest elements but that are very large, of the order of an electronvolt, for the heavy elements.

F. A. Jenkins; W. W. Watson

Fingerprint

Distinctive ridges that appear on the bulb of the inside of the end joint of each finger and thumb. These ridges have definite contours and appear in three general pattern types, the arch, the loop, and the whorl, each with general and specific variations of the pattern, dependent upon the shape and relationship of the ridges.

Patterns and Classification

Fingerprints are an infallible means of identification. In addition to their value in the apprehension of criminals, fingerprints can ensure personal identification for humanitarian reasons, such as in cases of amnesia, missing persons, or unknown deceased. Fingerprints are invaluable in effecting identifications in tragedies such as fire, flood, and vehicle crashes. In criminal matters, besides establishing the identity of the arrested person, fingerprint records provide a history of known offenders, or indicate when a person is a first offender. The vast majority of fingerprints maintained in the Criminal Justice Information Services Division of the Federal Bureau of Investigation (FBI) of the United States, the largest repository of fingerprints in the world, are civil records from military and other federal employees.

Patterns. Fingerprints fall into three general types of patterns. Each group bears the same general characteristics or family resemblance. The three general pattern types may be further divided into subgroups by means of smaller differences existing between the patterns in the same general group. The arch group includes the plain arch and the tented arch. The loop group includes the radial and ulnar loops. The whorl group includes four types of whorls, the plain whorl, central pocket loop, double loop, and accidental whorl.

Pattern area. The pattern area is that part of a fingerprint in which appear the cores, deltas, and ridges that are used for classification. The pattern areas of loops and whorls are enclosed by type lines, which may be defined as the two innermost ridges which start parallel, diverge, and surround or tend to surround the pattern area.

Focal points. Within the pattern areas of loops and whorls are the focal points which are used in the detailed classification. These points are called the delta and the core. The delta is that point on a ridge at or in front of and nearest the center of the divergence of the type lines. A core may be defined as that point on a ridge which is located in the approximate center of the finger impression.

Loop. The loop (**Fig. 1**) is that type of pattern in which one or more of the ridges enter on either side of the impression, recurve, touch or pass an imaginary line drawn from the delta to the core, and terminate or tend to terminate on or toward the same side of the impression from which such ridge or ridges entered. In the loop-type pattern the ridges intervening between the delta and core are counted.

The terms "radial" and "ulnar" are derived from the radius and ulna bones of the forearm. The classification of loops is determined by the way the loops flow on the hand. Loops which flow in the direction of the ulna bone, toward the little finger, are called ulnar loops, and those which flow in the direction of the radius bone are called radial loops.

Arches. In plain arches (**Fig. 2***a*) the ridges enter on one side of the impression and flow or tend to

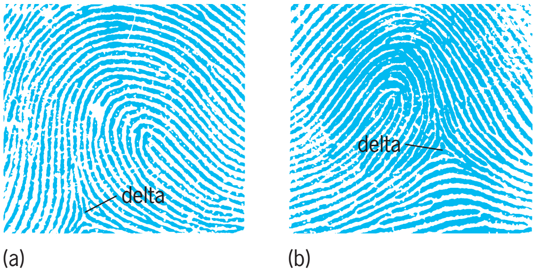
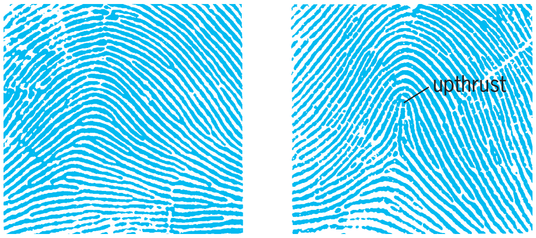


Fig. 1. **Loop fingerprint patterns. (*a*) Twelve-count ulnar loop, right hand. (*b*) Radial loop, right hand. (*Federal Bureau of Investigation*)**



(a) (b)

Fig. 2. Arch fingerprint patterns. (*a*) Plain arch. (*b*) Tented arch. (*Federal Bureau of Investigation*)

flow out the other side with a rise or wave in the center.

A tented arch (Fig. 2*b*) is that type of pattern which possesses either an angle, an upthrust, or two of the three basic characteristics of a loop. That is, the tented arch may lack a delta, ridge count, or recurve, which are characteristics of the loop.

Whorl. The whorl is that type of pattern in which at least two deltas are present with a recurve in front of each. This definition is very general. The pattern may be subdivided into four subgroups for extension purposes in large groups where whorls predominate.

Plain whorl. This pattern is the simplest and commonest of the whorl subdivisions. It has two deltas and at least one ridge making a complete circuit, which may be spiral, oval, circular, or any variant of a circle. An imaginary line drawn between the two deltas must touch or cross at least one of the recurving ridges within the inner pattern area.

Central pocket loop. This pattern is that type of whorl which has two deltas and at least one ridge making a complete circuit. An imaginary line drawn between the two deltas must not touch or cross any of the recurving ridges within the inner pattern area. In lieu of a recurve in front of the delta in the inner pattern area, an obstruction at right angles to the line of flow will suffice. The inner line of flow is determined by drawing an imaginary line between the inner delta and the center of the innermost recurve or looping ridge.

Double loop. This pattern consists of two separate loop formations and two deltas.

Accidental whorl. This is a pattern consisting of a combination of two "different" types of pattern, with the exception of the plain arch, with two or more deltas; or a pattern which possesses some of the requirements for two or more different types; or a pattern which conforms to none of the definitions (**Fig. 3**).

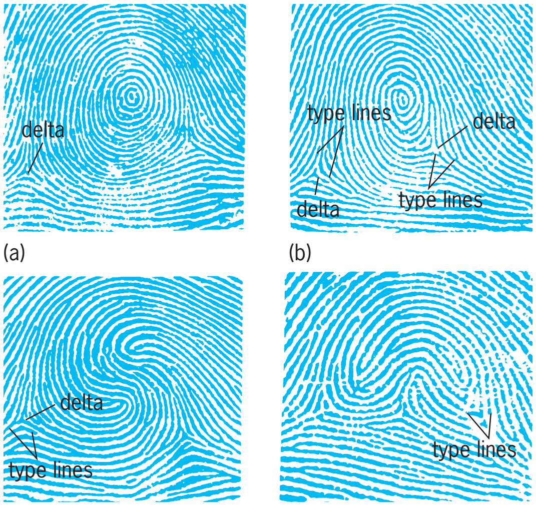
The technique of whorl tracing depends upon the establishment of the deltas. Every whorl has two or more deltas. When the deltas have been located, the ridge emanating from the lower side or point of the extreme left delta is traced until the point nearest or opposite the extreme right delta is reached. The number of ridges intervening between the tracing ridge and the right delta is then counted. If the ridge traced passes inside of (above) the right delta, and three or more ridges intervene between the tracing ridge and the delta, the tracing is designated as "inner" or I. If the ridge traced passes outside (below) the right delta, and three or more ridges intervene between the tracing ridge and the right delta, the tracing is designated as "outer" or O. All other tracings are designated as "meeting" or M. When the ridge traced ends abruptly, the tracing drops to the ridge immediately beneath and continues from there.

Classification. Although the system used in manual (noncomputerized) fingerprint files by the Federal Bureau of Investigation is based on the system devised by Edward Henry, it has been necessary to amplify and extend the Henry system with numerous additional subdivisions. The 10 fingers are considered as a unit in obtaining the complete classification. This classification is expressed in a formula consisting of a combination of letters and numerals. It reflects the general characteristics of the patterns in all fingers and permits filing in proper sequence for reading reference.

The fingerprint pattern types are indicated at the bottom of each finger block. Under the index fingers the appropriate capital letters are placed for every pattern except the ulnar loop, which is designated by a diagonal line slanting in the direction of the loop. Whorls in any finger are designated by the letter W. Under all other fingers, the appropriate small letter or diagonal line is placed for every pattern.

Formula. The classification formula is generally composed of the following divisions: primary, secondary, small-letter group, subsecondary, major, final, and key.

For the purpose of obtaining the primary classification, numerical values are arbitrarily assigned to each of the 10 finger blocks. Fingers numbers 1 and 2 have a value of 16, fingers numbers 3 and 4 have a value of 8, fingers numbers 5 and 6 have a value of 2, and fingers numbers 9 and 10 have a value of 1. Wherever a whorl appears, it assumes the value



(a) (b) (c) (d)

Fig. 3. Whorl fingerprint patterns. (*a*) Plain whorl, O tracing. (*b*) Central pocket loop, O tracing. (*c*) Double loop, I tracing. (*d*) Accidental whorl, I tracing; this is a whorl over a loop. (*Federal Bureau of Investigation*)

of the space in which it is found. The summation of the numerical values of the whorl type pattern, if any, appearing in fingers 1, 3, 5, 7, 9, plus an arbitrary 1, furnishes the denominator of the primary. The summation of the values of the whorls, if any, in fingers 2, 4, 6, 8, 10, plus an arbitrary 1, furnishes the numerator of the primary. For example, the primary classification on the set of fingerprints (**Fig. 4**) is derived from the location of the whorls in fingers 7, 9, and 10. The numerator is based on the whorl in the tenth finger, which has a value of 1. This plus the 1 arbitrarily added gives the 2 in the formula. The denominator depends upon the values of 2 and 1 given the whorls in fingers 7 and 9. This total of 3 plus the arbitrary 1 yields the 4 that is shown in the denominator.

It should be noted that, after the primary is obtained, the entire remaining portion of the classification formula is based upon the arrangements of the impressions appearing in the right hand as the numerator over the impressions appearing in the left hand as the denominator.

Secondary. The secondary classification appears to the right of the fractional numerals which represent the primary. It is shown in the formula by capital letters representing the types of patterns appearing in the index fingers of each hand.

Small letters. The small-letter group includes the prints with an arch or tented arch in any finger or a radial loop in any except the index fingers. Such small letters are brought up into the classification formula in their proper relative positions, either to the left or right of the secondary. A dash is used to indicate the absence of each small letter between the index fingers and another small letter, or between two small letters.

Subsecondary. The subsecondary classification is obtained by grouping the ridge counts of loops and the ridge tracings of whorls. The ridge counts are translated into small and large, represented by symbols I and O. The whorl tracings are brought up as 1, M, or O, denoting inner, meeting, or outer ridge tracings of the whorl types. Only six fingers are used to obtain the subsecondary, numbers 2, 3, 4, in the

APPLICANT

LEAVE THIS SPACE BLANK

Doe John M.

LAST NAME FIRST NAME MIDDLE NAME

SEX M

RACE W

SIGNATURE OF PERSON FINGERPRINTED

John M. Doe

CONTRIBUTOR AND ADDRESS

CHIEF
POLICE DEPARTMENT
CHARLOTTESVILLE, VA.

COMPANY AND ADDRESS

HT. (IN.) 70

WT. 168

DATE OF BIRTH 8/2/17

RESIDENCE OF PERSON FINGERPRINTED

1636 Rose Hill Dr.

HAIR Blk.

EYES Bl.

SIGNATURE OF OFFICIAL TAKING FINGERPRINTS

NUMBER

LEAVE THIS SPACE BL

TYPE OR PRINT ALL REQUESTED DATA

DATE FINGERPRINTED 4/16/59

PLACE OF BIRTH Crozet, Va.

CITIZENSHIP U.S.

CLASS. 20 | M | 2 | U | 000 | 20
| L | 4 | W | 000 |

REF.

SEE REVERSE SIDE FOR FURTHER

Fig. 4. Set of fingerprints showing classification and pattern types. (*Police Department, Charlottesville, Virginia*)

numerator, and 7, 8, and 9, in the denominator. The subsecondary is placed to the right of the secondary in the classification formula.

Major. The major classification is derived from the thumbs. When whorls appear, the major reflects the whorl tracings. When loops appear, however, a table of ridge count values is used to translate the ridge counts into small, medium, or large groups, designated by the letters S, M, L. The major is placed to the left of the primary in the classification formula.

Final. The final provides a definite order of filing the prints within the aforementioned subdivided groups. It is the ridge count of the loop in the right little finger. If the right little finger is not a loop, the ridge count of the left little finger is used. It is indicated at the extreme right of the classification formula.

Key. The key is obtained by counting the ridges of the first loop appearing on the fingerprint card, beginning with the right thumb, exclusive of the little fingers which are never considered for the key because they are reserved for the final. The key is always placed on the extreme left of the numerator of the classification formula.

File system. The use of the classification formula results in a file system susceptible to multiple expansion and segregates prints into groups on the basis of similar pattern characteristics in corresponding fingers. This enables the location of a record within minutes by examining a limited number of the millions of individual cards on file.

Latent prints. The Latent Fingerprint Section of the Federal Bureau of Investigation deals with the identification of single or latent (hidden) fingerprints developed at the scene of a crime or upon articles of evidence. This generally involves the examination of fragmentary latent finger, palm, or even foot impressions developed by appropriate processes on objects associated with various crimes.

The distinctive ridges of the palms of the hands and the soles of the feet closely resemble those of the fingerprint and can be identified in the same way. Footprints of persons having no fingers are taken for record purposes. Fingerprint, palm print, and sole print identifications have all been accepted as evidence by the courts. *See* EPIDERMAL RIDGES.

J. Edgar Hoover

Automated Fingerprint Identification Systems

An automated fingerprint identification system (AFIS) uses computer technology to collect, store, and identify friction skin ridge impressions from human fingers or palms.

Continuously evolving since the 1880s, fingerprint identification filing systems are used to identify one person's prints among a collection of record (known) fingerprints. Limited success in managing large hand-filed collections was achieved with the constant refinement and increase of classification subgroupings. For example, cards from persons with plain arch patterns on all ten fingers might be filed in a drawer, with further subgroupings according to five-year age groups and whether any of their fingerprint pattern areas include permanent scars.

With the introduction of computers into most large government record systems in the 1960s and 1970s, automated filing began evolving for the task of fingerprint identification in large collections. During the 1980s and 1990s, large fingerprint card repositories throughout most of the world began utilizing AFIS to greatly improve throughput for comparison and identification of fingerprint records. There was also an increase in accuracy for detecting duplicate records (previously undetected aliases). For latent prints collected from crime scenes or other evidence, AFIS enables identifications impossible in large manual (noncomputerized) finger and palm print databases. AFIS also links together unsolved crimes when the same (as yet unknown) person's prints are found. This creates additional leads for investigators. *See* DATABASE MANAGEMENT SYSTEM.

Automated palmprint identification system (APIS). These were introduced in the 1990s and operate similar to AFIS. The individualized nature of friction skin means that any ridge detail on the hands and feet can be encoded, searched, and identified in an automated system. *See* AUTOMATION.

Card reader. Scanners of various configurations are used to input record fingerprint cards into AFIS. Some automatic-feed-type card readers can input over 100 fingerprint cards per hour. Others are simple desktop models similar to those used with personal computers. *See* COMPUTER.

Consolidation. The discovery of two fingerprint records in a repository for the same person using different aliases is called a consolidation.

Encoding. To encode a latent print typically means to electronically mark (**Fig. 5**) the ridge details in preparation for launching a search against the fingerprint repository. Encoding record fingerprint cards often includes entering personal data about the person, verifying computer evaluation of legible friction ridge detail areas of prints, and verifying computerized classification of the pattern types for each finger.

Integrated Automated Fingerprint Identification System (IAFIS). This is the name of the FBI's nationwide AFIS repository located in Clarksburg, West Virginia. IAFIS links all 50 states and all federal agencies.

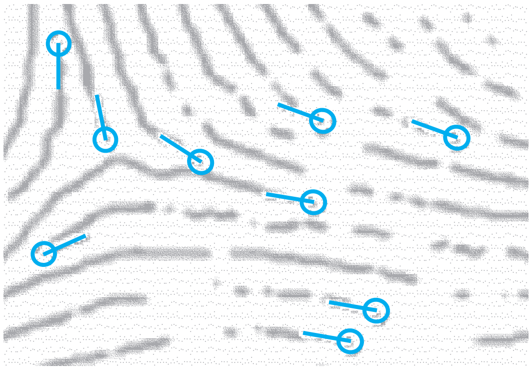


Fig. 5. An AFIS is electronically marked for plotting location (x/y axis) and orientation (rotation).

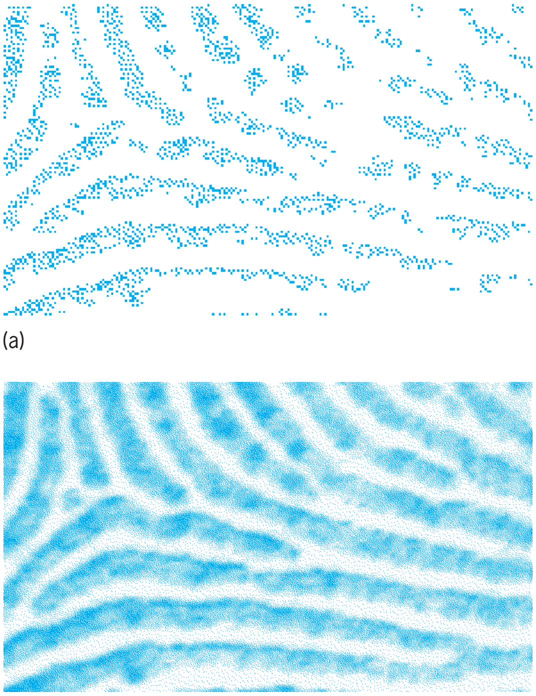
Latent print searches. All chance or accidental evidentiary finger or palm prints left on an item of evidence or at the scene of a crime are referred to as latent prints. This includes impressions in paint, ink, or blood, as well as those truly invisible before chemical development. About one-third of the latent fingerprints found at most crime scenes are suitable for AFIS searching. The remainder are often from lower finger joints, finger tips, sides of fingers, or other areas not normally found in an AFIS.

Latent workstation. This is usually a personal computer networked with an AFIS, and having latent print capture, encoding, searching, and comparison capabilities. It often includes a digital camera or a flatbed scanner for image capture.

Live scan. These tabletop or stand-alone devices for scanning friction skin ridges capture electronic finger or palm prints by using light or ultrasound.

Record print. All types of known, purposely recorded finger or palm prints are referred to as record prints. Live scan, inkless fingerprint cards, and powder-developed impressions lifted from skin are all record prints.

Matching score. AFIS compares fingerprints using mathematical matching algorithms, resulting in a matching score. For record print comparisons, an identification threshold may be set for the matching score in some systems. All matching scores above the threshold will be positive identifications with zero errors. Matching scores below the threshold may also be identifications as well, but require human analysis, comparison, evaluation, and verification to complete the identification process. Latent prints (**Fig. 6**) normally have insufficient area or clarity to enable automatic identification using a matching score.



(a)

(b)

Fig. 6. For comparison purposes, (*a*) evidence prints are seldom as legible as (*b*) prints on file.

Point. Ridge endings, bifurcations, and dots are all routinely called points, characteristics, Galton details, minutiae, or features.

Skeleton. Fingerprint images are sometimes thinned or skeletonized during automated feature extraction, enabling easier automated computer detection and encoding.

Smart card. Personal identification cards containing automated fingerprint-feature encoding of the authorized user are one version of a biometric smart card widely used. The user inserts the card into a card reader and places the appropriate finger on a live scan device. A fast and simple one-to-one verification is then completed, using the stored minutiae and the person's finger live scan image. Smart cards are used for automatic teller machine (ATM) cards for banks, access to military restricted areas or restricted computers, and a variety of private and government access control programs. Use of both a fingerprint and a personal identification number (PIN) is required with some smart cards.

Tracing. Tracing ridge detail on translucent plastic over latent print photographic enlargements was required for many early AFIS systems. The resulting ridge detail was similar to the skeleton sometimes produced during AFIS ridge point encoding, and lacked the typical background interference present on evidence. The hand-produced tracings were then reduced to life size and "read" into the AFIS through a camera or scanner. Modern AFIS procedures do not require tracing. However, it is still practiced by many agencies which believe it is faster and more accurate insofar as isolating valid ridge details versus electronically marking points on a computer monitor.

Finger, palm, and sole impression comparison and identification is based on three levels of friction ridge details: Level I—ridge flow or general pattern configuration (**Fig. 7**); Level II—ridge endings, bifurcations, and dots (**Fig. 8**); and Level III—ridge shape, including the shape of ridge edges and pores (**Fig. 9**). Level I information can be captured by sampling images at 150 pixels per inch (ppi). Level II detail needs 1000 ppi or higher resolution for best observation.

Automated fingerprint filing attempts in the early 1960s were computerized search programs which reported the groups of fingerprint cards to be searched based on Level I pattern and other class or group characteristics such as age, sex, race, and geographic location. Some systems, such as Miracode, linked together microfilm rolls of fingerprint images and computer records. These were the first fingerprint comparison workstations. None were efficient, and intensive human evaluation, classification, comparison, and verification were constantly involved.

Through collaborative efforts between the FBI, Rockwell International, and other laboratories in the United States, France, and the United Kingdom, the first AFIS using Level I and II details for searching and comparison was developed between 1967 and 1972. The Level II information recorded included the x and

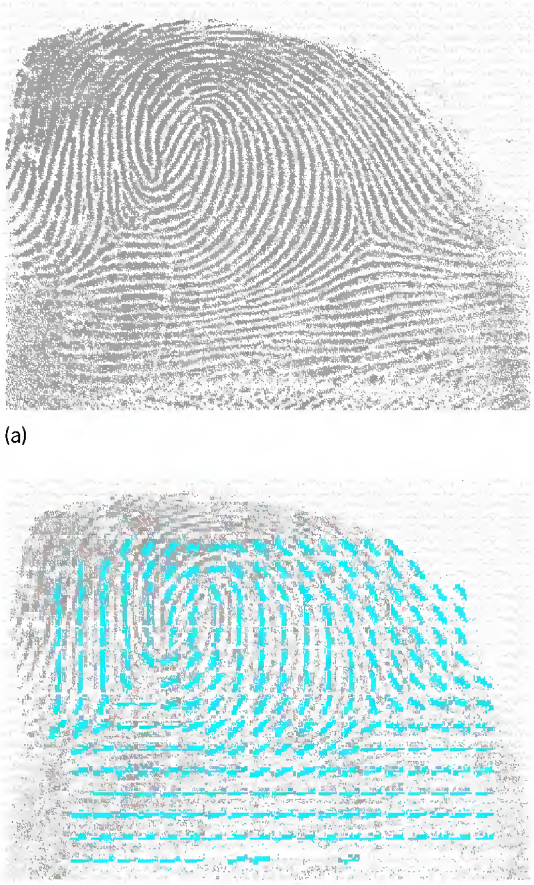


Fig. 7. Typical (*a*) inked fingerprint with (*b*) colored lines on the fingerprint representing Level I detail.

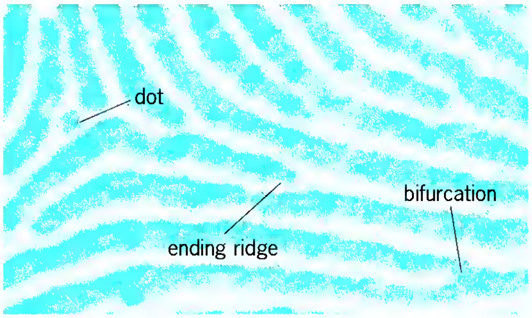


Fig. 8. Enlarged fingerprint showing Level II detail.

y coordinate axis, as well as the rotation or orientation of the ridge ending and bifurcation. Most AFIS ignore dots.

Smaller and more advanced AFIS have been developed in the decades since. Most use not only the x/y axis and rotation, but also intervening ridge counts between individual points and the nearest 3–8 neighboring points. Accuracy for automated comparison between inked prints have been enhanced to the point that in some systems a matching score threshold can be set to enable completely automated comparison and "full" identification with no human intervention and verification and zero error rate.

AFIS operations typically are divided between fingerprint repository (known print) and crime laboratory (questioned print) activities. All record or known prints are first checked by name and date of birth in an attempt to verify identification without wasting AFIS feature computer matching resources. If there is no personal-data-based identification, a technical search using fingerprint minutiae is launched in an attempt to detect existing fingerprint records under an alias name. In a background operating mode, newly arriving fingerprint cards each day are also compared against all unsolved-case latent prints in the AFIS. Crime laboratory experts proactively launch latent print searches against the record fingerprints, and also against other unsolved-case latent prints in an attempt to link related cases.

The National Crime Information Center Year 2000 (NCIC2000) is an FBI service that enables near-instant wanted-person checks using a single fingerprint scanned on a portable live scan unit carried in a patrol car. For example, a person's right index fingerprint can be transmitted electronically to the FBI, searched against a database of 250,000 wanted persons, and an arrest or not-wanted response returned within 2 min. Miniaturization even enables most cities' entire fingerprint repositories to be stored and managed in a personal computer with full AFIS crime-solving capabilities.

Cooperation between police agencies in the 1980s and 1990s resulted in the establishment of limited AFIS definition standards through the American National Standards Institute/National Institute of Standards and Technology (ANSI/NIST). As a result, most record or known fingerprints are captured, stored, and transmitted as 500 or 1000 pixels per inch (ppi) grayscale images using wavelet scalar quantization (WSQ) compression. Most evidence or latent prints are captured as 1000 ppi (or higher) grayscale or color images with no compression; that is, bitmap or tagged image format files are used. Sampling rates of less than 1000 ppi do not capture sufficient Level

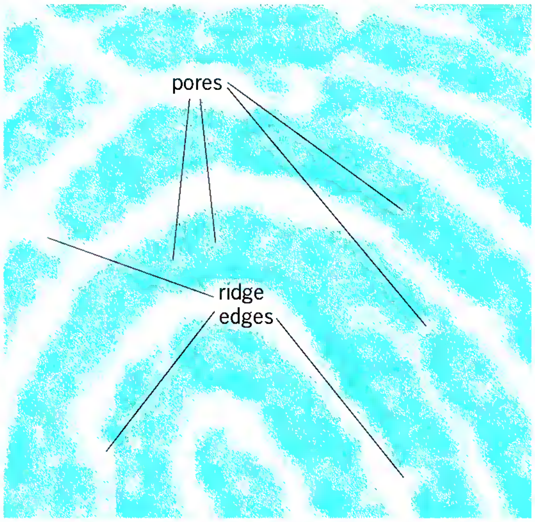


Fig. 9. Enlarged fingerprint showing Level III detail.

III detail to enable meaningful comparison with difficult evidence prints. Some agencies with 500 ppi card scan systems retain the paper fingerprint cards solely to enable latent print comparisons. *See* IMAGE PROCESSING.

Edward R. German

Bibliography. D. R. Ashbaugh, *Quantitative-Qualitative Friciton Ridge Analysis*, CRC Press, Boca Raton, FL, 1999; D. Caudle, Personal communication, The Phoenix Group, Pittsburg, KS, 1999; H. Lee and R. Gaensslen, *Advances in Fingerprint Technology*, CRC Press, Boca Raton, FL, 1994; R. Olsen, *Scott's Fingerprint Mechanics*, Thomas Publishing, Springfield, IL, 1978.

Finite element method

A numerical analysis technique for obtaining approximate solutions to many types of engineering problems. The need for numerical methods arises from the fact that for most engineering problems analytical solutions do not exist. While the governing equations and boundary conditions can usually be written for these problems, difficulties introduced by either irregular geometry or other discontinuities render the problems intractable analytically. To obtain a solution, the engineer must make simplifying assumptions reducing the problem to one that can be solved, or a numerical procedure must be used. In an analytic solution, the unknown quantity is given by a mathematical function valid at an infinite number of locations in the region under study, while numerical methods provide approximate values of the unknown quantity only at discrete points in the region. In the finite element method, the region of interest is divided into numerous connected subregions or elements within which approximating functions (usually polynomials) are used to represent the unknown quantity.

Development. The physical concept on which the finite element method is based has its origins in the theory of structures. The idea of building up a structure by fitting together a number of structural elements is quite a natural one (**Fig. 1**) and was used in the truss and framework analysis approaches employed in the design of buildings and bridges in the early 1900s. By knowing the characteristics of individual structural elements and combining them, the governing equations for the entire structure could be obtained. This process produces a set of simultaneous algebraic equations. The limitation on the number of equations that could be solved imposed a severe restriction on the analysis. *See* STRUCTURAL ANALYSIS.

The finite element method evolved into a practical analysis tool as a result of development work in aircraft structural analysis. This was due to demanding analysis requirements coupled with several technological advances in the 1950s and 1960s. The first of these was the formulation of structural problems in terms of matrix algebra, thus forming the governing algebraic equations in a routine manner. Next, the introduction of the digital computer made possible the solution of large-order systems of equations. Finally, the theoretical concepts of finite element analysis crystallized during the same period. *See* DIGITAL COMPUTER; LINEAR ALGEBRA; MATRIX THEORY.

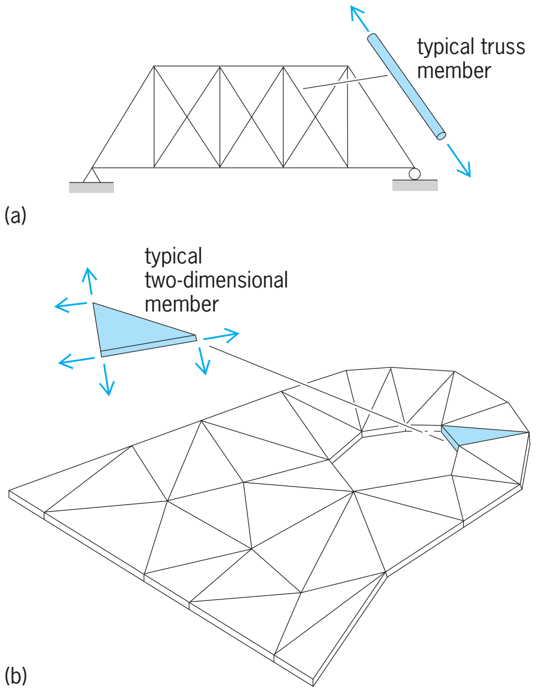


Fig. 1. Structures modeled by fitting together structural elements: (*a*) truss structure; (*b*) two-dimensional planar structure.

Development work since then has caused the finite element method to emerge as one of the most powerful approaches for obtaining approximate solutions to a wide range of problems in mathematical physics. The method has achieved acceptance in nearly every branch of engineering and is the preferred approach in structural mechanics and heat transfer. Its application has extended to soil mechanics, fluid flow, magnetic field calculations, and other areas. *See* HEAT TRANSFER; SOIL MECHANICS.

Mathematical methods. Due to the broadened scope of the finite element approach and the recognition that it is one of the most powerful methods for the approximate solution of boundary value problems, it has attracted the attention of mathematicians who have endeavored to establish a solid mathematical basis for the method. It is, in essence, a variational method of approximation, employing the Rayleigh-Ritz-Galerkin approach. A given region is represented as a collection of a number of geometrically simple regions (finite elements) connected at nodal points. Simple mathematical functions, generally polynomials, are chosen for each element. *See* CALCULUS OF VARIATIONS.

Conventional finite element approaches use Lagrange polynomials as displacement shape functions. An alternate approach, the P-finite element method, uses Legendre polynomials. In the P method, more sophisticated elements are generated by using high-order Legendre polynomials to produce “elasticity”-type solutions for a wide range of problems. *See* HYPERGEOMETRIC FUNCTIONS; LEGENDRE FUNCTIONS; ORTHOGONAL POLYNOMIALS.

Another advance in finite element technology is the resurgence of integral methods, which has led to the development of the boundary element method into an accepted engineering analysis tool. Only the boundary of the region is discretized into individual elements. For two-dimensional problems the elements are one-dimensional lines on the boundary, while for three-dimensional problems the elements are two-dimensional surface elements. In contrast to conventional finite element approaches, there are no nodes within a region. This has two advantages: (1) Only the boundaries need to be defined, so that less input is required, and (2) a smaller set of equations needs to be solved. Moreover, the boundary element method, like the P method, produces high-accuracy solutions. *See* INTEGRAL EQUATION.

Element formulation and types. The formulation of the individual finite element matrices is central to the method, since the total solution to a problem is constructed of linear combinations of the individual element behavior. In structural mechanics, the unknown field variables of displacements or stresses are defined in terms of values at the node points, which are the unknowns of the problem.

The accuracy of the solution depends not only on the number and size of the finite elements used but also on the functions selected for the elements. These functions must be carefully chosen and must satisfy specific criteria to ensure convergence to the “correct” solution.

Using these functions, a number of approaches can be used to formulate the element matrices (stiffness, stress, induced strain, mass, and so on). The first of these is called the direct approach because physical reasoning is used to establish the element matrices. The other approaches are more theoretical and versatile, allowing the method to be extended to other fields. Here the calculus of variations is used to derive energy principles, which are then employed in deriving the element formulations.

In solid mechanics problems, the element shape functions represent a displacement field within each element (commonly called the displacement method). They could also be chosen to represent stresses (the stress or equilibrium method) or a combination of displacements and stresses (the hybrid method). For most problems in solid mechanics, the displacement method is the easiest to apply, and consequently it is the most widely used approach.

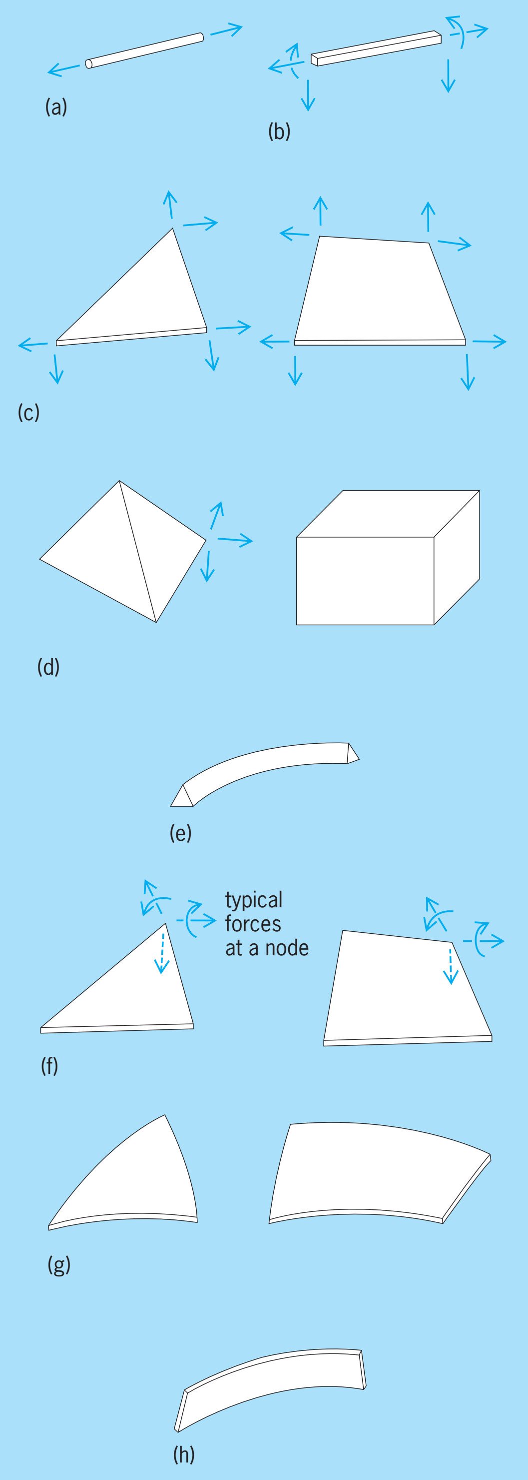


Fig. 2. Element types: (*a*) bar; (*b*) beam; (*c*) membrane elements; (*d*) solid elements; (*e*) axisymmetric solid; (*f*) plate-bending elements; (*g*) curved-shell elements; (*h*) axisymmetric thin-shell element.

Some of the more commonly used finite elements are depicted in **Fig. 2**. The bar (Fig. 2*a*) is the simplest type of element and can resist only axial loads. These elements can be used to model truss-type structures. The beam element (Fig. 2*b*) is similar to the bar element, but can also take moments and shear loads necessary for representing frame structures. The membrane elements (Fig. 2*c*), grouped together, are basic in finite element analysis of two-dimensional problems. They are thin-membrane elements, triangular and quadrilateral in shape, that are loaded only in their own plane.

Extension of the concept to three dimensions results in a family of solid elements (Fig. 2*d*), of which the tetrahedron and hexahedron are the most common shapes. These elements are used for the solution of three-dimensional stress-analysis problems. Another type of solid element is the solid of revolution (Fig. 2*e*). The remaining elements in Fig. 2 are classified as bending elements in that they can also take loads normal to their surfaces. The plate-bending elements (Fig. 2*f*) are used to model flat plates and shells or thin-walled members. A more accurate representation of a shell structure can be obtained by using curved shell elements (Fig. 2*g*). Finally, the shell-of-revolution element (Fig. 2*h*) is applicable to a restricted class of thin-shell problems. Elements with additional nodes placed between the corner nodes, using higher-order functions and providing more accurate results, are becoming more accepted.

Application and implementation procedure. The complexity and size of problems which can be addressed by the finite element method has increased. This has been due to advances in element formulation and in solution algorithms, as well as in computer software and hardware technology. The use of the method is illustrated by a representative application taken from the aerospace industry.

Figure 3 shows a finite element model of a business jet which is typical of modeling in aircraft structures. Both the fuselage and wing construction consist of thin metal sheet wrapped around a framework. Bending effects are small enough to be neglected in the skin, which is modeled with membrane elements (Fig. 2*c*). The framework structure (spars and ribs in the wing, frames and longerons in the fuselage) is modeled with bars and beams (Fig. 2*a* and *b*) as well as membrane and bending plate elements (Fig. 2*c* and *f*). The model shown is for a medium-complexity business jet; larger aircraft, such as the Boeing 747, require more complex models. The finite element analysis of the 747 wing-body region alone required the solution of over 7000 equations. Problems requiring the solution of hundreds of thousands of equations are now common. *See* AIRFRAME; FUSELAGE; WING STRUCTURE.

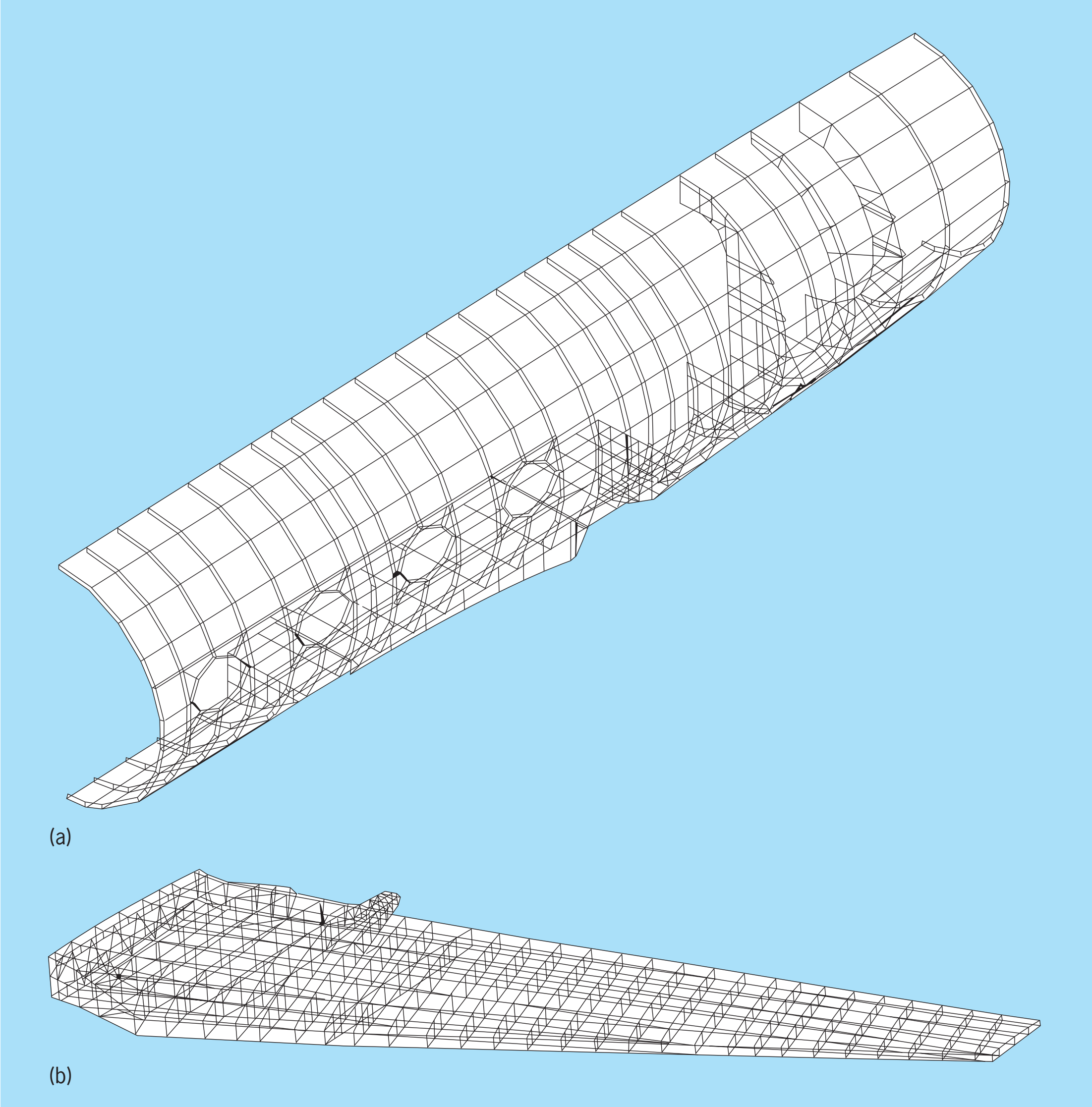


Fig. 3. Finite element method of (*a*) fuselage section and (*b*) wing of a business jet.

While aircraft companies were the first to use the finite element method extensively, today all major designers and manufacturers of automobiles, ships, power plants, bridges, and computers use this analytical tool. The finite element method enables the automobile industry to produce vehicles that are lighter and thus more efficient (using structural optimization techniques) as well as safer (using crash worthiness analysis with the method). In addition to applications in static structural analysis, the method is used to compute vibration modes and perform dynamic analysis as well as to calculate buckling loads for failure analysis.

Finite element analysis is most often implemented with large general-purpose programs capable of handling many types of problems. These programs are very costly to develop and are often key components of computer-based design systems which incorporate interactive graphics and data management systems. The development and maintenance of these programs has shifted from companies that are primarily users today, to companies that specialize in the development of these enhanced products. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

The tasks necessary to perform a finite element analysis are the following:

1. *Model generation*. The first step requires a great deal of modeling expertise. The level of complexity required to meet the intent of the analysis must be established, as well as the types of elements to be employed. Input data describing the coordinates of the grid points and the element topology and properties, as well as boundary conditions and load data, must then be generated. Today these tasks are done automatically in preprocessing modules that are part of the commercial products. In addition, improved interfaces with computer-aided design systems allowing automatic access to wire-frame and solid geometry are moving toward an integrated computer-aided engineering environment. *See* COMPUTER-AIDED ENGINEERING.

2. *Data verification*. The large amounts of data must be checked, since the consequences of error are costly. This is accomplished with graphics modules and special data checking routines as well as error guidelines which are contained in the programs.

3. *Analysis*. The governing matrix equations for the finite element model are generated and solved. For static structural analysis problems, a set of linear algebraic equations are solved, yielding the deflections at the grid points. Substitution into an additional set of equations gives the stresses or forces in the individual elements.

4. *Postprocessing*. The output of the analysis is organized and edited to produce efficient-to-use formats. This includes conversion to report formats, and extensive use of computer-generated graphics showing internal loads and deflections.

The finite element method has found a tremendous number of applications. This is due to its generality in analyzing almost any type of structure without any restrictions on material properties or boundary conditions. In addition, it is readily adaptable to the design process, wherein crude models are employed in the conceptual stages, and more detailed models as the design evolves. Moreover, it is used to reduce the cost of test programs by using analytical models to minimize the number of tests required. Significant advances, including crack propagation capabilities, shape optimization as an outgrowth of structural optimization, the boundary element method, and the P method incorporating automatic error estimation, have further enhanced the scope and the use of the method. Theodore Balderes

Bibliography. K.-J. Bathe, *Finite Element Procedures*, 2d ed., 1995; C. A. Brebbia and J. Dominguez, *Boundary Elements: An Introductory Course*, 2d ed., 1992; R. D. Cook, D. S. Malkus, and M. E. Plesha, *Concepts and Applications of Finite Element Analysis*, 3d ed., 1989; R. H. MacNeal, *Finite Elements: Their Design and Performance*, 1994; Ch. Schwab, *p- and hp-Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics*, 1999; O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method*, 2 vols., 4th ed., 1994.

Finite mathematics

These parts of mathematics which deal with finite sets. More generally, finite mathematics is taken to include those fields which make no use of the concept of limit; thus linear programming, which finds maxima and minima through the four arithmetic operations, is included, whereas classical optimization, which uses differential calculus, is excluded. *See* LINEAR PROGRAMMING; OPTIMIZATION.

In general, finite mathematics has grown in popularity because it represents, in a sense, the maximum that can be taught without introducing the difficult concepts of limits; thus it can serve as a valuable college course for nonscience majors. There is, however, only the above criterion to unify the field. Brief descriptions will be given of several of the subjects generally included in finite mathematics.

Set theory. An algebra of finite sets can be developed by introducing the operations of union, intersection, and difference. This is necessarily a very elementary set theory, and the usual paradoxes of infinite sets are avoided. *See* BOOLEAN ALGEBRA; SET THEORY.

Logic. As with sets, an algebra of propositions (the propositional calculus) can be developed with the operations of conjunction, disjunction, negation, and implication. There is a close parallelism between this algebra and the set algebra discussed above. Thus if P and Q are the sets of those x's for which the statements $p(x)$ and $q(x)$ respectively are true, then the intersection $P \cap Q$ is the set of x's for which the conjunction $p(x) \wedge q(x)$ is true. *See* LOGIC.

Algebraic structures. Finite groups, rings, and fields are interesting both in themselves and because of their applications. Thus, chemists are interested in finite groups as representing the possible symmetries of a molecular structure, and statisticians might study finite fields for experimental design purposes. *See* COMBINATORIAL THEORY; CRYSTAL STRUCTURE; GROUP THEORY; RING THEORY.

Linear algebra. Matrices and finite-dimensional vectors represent linear systems; they are thus of great interest, for example, in economics, in which linearity assumptions are at least approximately valid. Thus, if additivity and linear returns to scale are assumed, some economic systems can be represented by a technological matrix $A = (a_{ij})$, where a_{ij} is the amount of good j needed in the production of one unit of good i. The input and output of the system, vectors x and y respectively, are then related by the equation $x = Ay$. *See* LINEAR ALGEBRA; MATRIX THEORY.

Probability. Finite probability spaces and the related random variables are of interest. Most concepts in probability theory—except the limit theorems—can be explained within the context of finite mathematics. In particular, finite Markov chains can be studied by the introduction of matrices. *See* PROBABILITY; STOCHASTIC PROCESS.

Linear programming. An important application of linear algebra arises in economics if it is assumed, as above, that input and output are related by a linear equation, and that profits are also a linear function of output, of the form $P = c^t y$ (where c is also a vector, and the superscript t indicates that it is transposed, to form a row vector). Then, if a maximum input vector b is permitted, the economic problem is to find an output (or activity) vector y which will maximize $c^t y$, subject to the constraints $Ay \leq b$, $y \geq 0$ (because all outputs or activities must be nonnegative). The generic problem of maximizing (or minimizing) a linear function subject to linear (inequality or equation) constraints is known as a linear program. It can be proved, first, that the set of feasible points (those satisfying the constraints) is a polyhedral subset of space, and has therefore a finite number of extreme points (vertices); and second, that the maximum or minimum (if it exists) can always be found at one of these extreme points. Methods then exist that permit a systematic examination of these extreme points to determine the optimum.

Game theory. In the theory of finite two-person games, an $m \times n$ matrix $A = (a_{ij})$ is given. Players one and two choose respectively a row i and a column j of the matrix; player two then pays player one the amount (payoff) a_{ij} which appears in the ith row and jth column of the matrix. An entry a_{kl} is known as a saddle point if it is simultaneously the largest element in its column and the smallest in its row, and it is generally recommended that the players choose the row and column which give the saddle point, if one exists. Unfortunately, games do not generally have saddle points; this has led to the concept of a mixed strategy, which is best defined as a probability distribution over the rows (for player one) or columns (for player two) of the matrix. Mixed strategies take the form of vectors $x = (x_1, \ldots, x_n$ and $y = (y_1, \ldots, y_n)$ respectively, and give rise to an expected payoff equal to the matrix product xAy^t. It can then be shown that, with this payoff, a game always has a saddle point in mixed strategies. Linear programming can be used to compute the strategies (known as optimal strategies) which give rise to the saddle point. *See* GAME THEORY.

Dynamic programming. Certain physical systems can be represented by multistage processes for which, at each stage, the output is a function of that at the preceding stage (the state), an action taken by a decision maker (the control), and possibly a random event. An example occurs in inventory control, in which the inventory of a given item depends on the previous day's inventory (the state), the amount produced (the control), and the external demand (random event). This gives rise to a constraint space of very high dimension. The technique of dynamic programming reduces this sequence to one of one-dimensional problems, which are generally much simpler to solve.

Networks. A network is a collection of points (nodes), together with arcs joining certain pairs of nodes which may represent communications links between various locations, roads joining certain cities, and so on. Numbers representing, for example, the capacity or length of the arcs are frequently assigned. Problems considered may be of a combinatorial nature (finding the number of possible connections between two nodes) or of an optimization nature (finding the shortest path, or maximizing the amounts shipped between two nodes).

Guillermo Owen

Bibliography. S. J. Farlow, *Finite Mathematics and Its Applications*, 2d ed., 1994; J. G. Kemeny et al., *Introduction to Finite Mathematics*, 3d ed., 1974; D. P. Maki and M. Thompson, *Finite Mathematics*, 4th ed., 1995; A. Mizrahi and M. Sullivan, *Finite Mathematics*: *An Applied Approach*, 7th ed., 1995.

Fiord

A segment of a troughlike glaciated valley partly filled by an arm of the sea. It differs from other glaciated valleys only in the fact of submergence. The floors of many fiords are elongate basins excavated in bedrock, and in consequence are shallower at the fiord mouths than in the inland direction. The seaward rims of such basins represent lessening of glacial erosion at the coastline, where the former glacier ceased to be confined by valley walls and could spread laterally. Some rims are heightened by glacial drift deposited upon them in the form of an end moraine.

Fiords occur conspicuously in British Columbia and southern Alaska (**Figs. 1** and **2**), Greenland, Arctic islands, Norway, Chile, New Zealand, and Antarctica—all of which are areas of rock resistant to erosion, with deep valleys, and with strong former glaciation. Depths of the floors of many fiords exceed 2600 ft (800 m); the deepest on record is 6379 ft (1933 m), in Skelton Inlet, Antarctica. Depths are the result of enhanced glacial erosion caused by the rapid flow due to confinement of the glacier between valley sides.

Fig. 1. Fiord in Alaska. (*W. W. Atwood*, *USGS*)

Fig. 2. Typical fiord-side scenery of southeastern Alaska in June. Steep forested slopes give way upward to rocky ice-scoured upland. (*R. M. Chapman*, *USGS*)

Some shallow fiords were glaciated while their floors stood above sea level, and were submerged later. Others were deepened below the level of the sea at the time of glaciation. A valley glacier flowing into seawater continues to exert stress on its floor until flotation occurs. A vigorous glacier of density 0.9 and thickness of 3300 ft (1000 m) would continue to erode its floor even when submerged to a depth of nearly 3000 ft (900 m). As net submergence of coasts through postglacial rise of sea level is unlikely to have much exceeded 300 ft (100 m), excavation below the sea level of the time probably occurred in most valleys that are now deep fiords. *See* GLACIATED TERRAIN; OCEAN CIRCULATION.

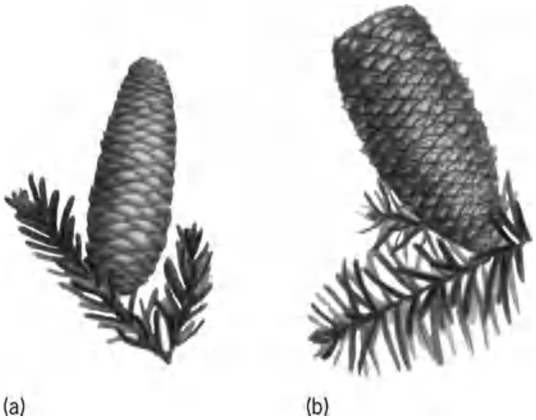
R. F. Flint

Bibliography. R. F. Flint, *Glacial and Quaternary Geology*, 1971; H. J. Freedland et al. (eds.), *Fjord Oceanography*, 1980; J. W. Gregory, *The Nature and Origin of Fiords*, 1913; J. P. Syvitski et al., *Fjords*, 1986.

Fir

Any tree of the genus *Abies*, of the pine family, characterized by erect cones, by the absence of resin canals in the wood but with many in the bark, and by flattened needlelike leaves which lack definite stalks. The leaves usually have two white lines on the underside and leave a circular scar when they fall. The native fir of the northeastern United States and adjacent Canada is *A. balsamea*, which attains a height of 75 ft (23 m) and has resinous buds. Its principal uses are for paper pulp, lumber, boxes, and crates, and as a source of the liquid resin called Canada balsam. In the eastern United States the fir is commonly used as a Christmas tree. It does not do well as an ornamental tree in areas where the summers are hot. *See* PINALES.

The Fraser fir (*A. fraseri*) is a similar species found in the southern Appalachians. In this species the bracts grow out beyond the cone scales (see **illus.**).



Cones and branches of (*a*) balsam fir (*Abies balsamea*) and (*b*) Fraser fir (*A. fraseri*).

Several species of *Abies* grow in the Rocky Mountains region and westward to the Pacific Coast. The most important is the white fir (*A. concolor*), also known as silver fir, a large tree which may reach a height of 120 ft (36 m) and is characterized by blue-green needles 2–3 in. (5–7.5 cm) long. It is often cultivated as an ornamental in the eastern United States, and also furnishes lumber for construction timber, boxes, crates and general millwork.

Other western species of commercial importance are the subalpine fir (*A. lasiocarpa*), grand fir (*A. grandis*), Pacific silver fir (*A. amabilis*), California red fir (*A. magnifica*), and noble fir (*A. procera*). *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Fire

A rapid but persistent chemical reaction accompanied by the emission of light and heat. The reaction is self-sustaining, unless extinguished, to the extent that it continues until the fuel concentration falls below a minimum value. Most commonly, it results from a rapid exothermic combination with oxygen by some combustible material. Flame and heat also may result from a reaction involving an agent other than oxygen. Thus, certain reactive metals such as zinc will burn in an atmosphere of chlorine.

Flame, the visible manifestation of fire, results from a heating to incandescence of minute particulate matter composed principally of incompletely burned fuel. The color of the flame depends upon the material undergoing reaction and the temperature. *See* COMBUSTION; FLAME; FOSSIL FUEL.

Francis J. Johnston

Fire detector

A device designed to initiate a desired action in response to a change in its surroundings caused by an unwanted fire. Heat detectors respond when the surrounding temperature exceeds a specific minimum value or when the rate of rise exceeds a specified rate. Heat-detector sensing elements are normally mechanical, employing components which melt or expand when heated, but can also be electronic, using elements which produce thermoelectric voltages or change in conductivity.

Smoke detectors are also commonly used fire detectors. They respond when the concentration of solid or liquid particles in the air exceeds a selected value. In ionization-type smoke detectors, a small radioactive source is employed to make the air within the sensing chamber slightly electrically conductive. The presence of particles reduces this conductivity in proportion to the number of particles present. In photoelectric-type smoke detectors, a light source and light-sensitive receiver are arranged so that light from the source does not normally strike the receiver. Smoke particles reflect light from the source onto the receiver in proportion to the number of particles present.

Flame detectors contain receiving elements sensitive to light in certain specific infrared or ultraviolet wavelengths characteristic of flames. Flame detectors can detect a fire within a few milliseconds and are therefore most often used to activate explosion-suppression systems.

The actions initiated by fire detectors include sounding of audible and visual alarms, signal transmission to remote monitoring stations, activation of extinguishing systems, and emergency shutdown of equipment or processes which might increase the severity of the fire hazard. *See* AUTOMATIC SPRINKLER SYSTEM; FIRE EXTINGUISHER.

Richard W. Bukowski

Bibliography. J. L. Bryan, *Fire Suppression and Detection Systems*, 3d ed., 1993; National Fire Protection Association, *Fire Protection Handbook*, 18th ed., 1997.

Fire extinguisher

Fire may be extinguished by the following methods: (1) cooling the burning materials; (2) blanketing the fire with inert gas that chokes it for lack of oxygen; (3) introducing materials that inhibit combustion; and (4) covering the burning matter with a blanket or a layer of solid particles that prevent access of air. Fire extinguishers operate on one or a combination of these principles.

The method or methods that are used to put out a fire depend upon its nature. Burning magnesium reacts with water according to the equation: $Mg + H_2O \rightarrow Mg + H_2$. This reaction is itself a form of combustion which is supported by water instead of oxygen. In fact, the hydrogen evolved will burn in the air and possibly create a worse situation. However, to pursue this example still further, the action of water on a magnesium fire will depend not only on these chemical factors but also on the relation between the amount of water and the amount of burning magnesium (the surface area of magnesium covered by flame). A small amount of water on a large magnesium fire would make matters worse; but a very large amount of water on a small magnesium fire would quench it because of the cooling effect. Therefore, both the chemical and the physical nature of the fuel and the size of the fire have an important bearing on the effectiveness of a given extinguishing technique.

Water. Water is the most effective cooling agent used in fire extinguishing. The generation of steam also drives away the air and forms a blanket, but being less dense than air, it is rapidly displaced.

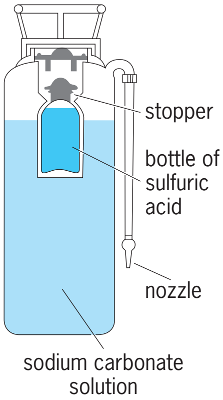
Wetting agents and foaming agents increase the effectiveness of water. In the case of burning cotton bales, for example, the addition of a wetting agent will cause the water to penetrate the burning material and extinguish the fire more efficiently. Foaming agents make it possible to produce a heavy blanket of foam over the fire. This excludes air and cools the burning material at the same time. Attachments that feed wetting and foaming agents into the fire hose stream are now common.

The use of water in the form of fog rather than as a spray or stream is very effective under some conditions. In fighting an oil fire, for example, fog is effective in cooling the fire and providing a blanket, whereas a stream of water would serve only to disperse the oil and spread the fire. The greatest benefit of fog is that it provides a protecting screen and allows the fire fighter to work in close to the fire.

In the small, first-aid, water fire extinguishers, a propellant must be provided. Usually this is carbon dioxide, which is either generated when needed (see **illus.**) or stored in a cartridge. When a foaming agent is used in such an extinguisher, the foam is rendered still more effective by the inert gas that fills the bubbles.

Antifreeze agents must be added if the water extinguisher is placed outdoors in cold climates. Commonly, calcium chloride is used, but in certain cases (the so-called loaded stream extinguishers), the antifreeze salts may be such as to have a definite fire retarding effect.

There are many fires, as pointed out above, on which water should not be used. On oil fires, for example, water is effective only as a foam or in some cases as a fog; a water stream should never be used. On electrical fires, the use of water can prove extremely dangerous, and in the case of electric generators and motors, it is completely ruinous. Fog nozzles



Soda-acid fire extinguisher. When it is inverted, the stopper falls out of place, allowing sulfuric acid to react with sodium carbonate to produce carbon dioxide.

can sometimes be used effectively, but there is always an element of added danger to the fireman whenever water is used near an electrical fire. Carbon dioxide or dry extingishers are decidedly preferred.

Automatic water sprinkler systems are now the most common form of fire protection in industrial plants and large buildings. The installation of these systems is the greatest single factor to be credited for the sharply reduced incidence of disastrous fires. *See* AUTOMATIC SPRINKLER SYSTEM.

Carbon dioxide. A safe and effective extinguisher for all confined fires, carbon dioxide, CO_2, acts as an inert blanket, and because it is heavier than air, it will exclude oxygen very efficiently from a fire on the floor of a building or in a vat or similar vessel. It is not effective in an elevated location or outdoors where the wind can blow the gas away.

Carbon dioxide is supplied as liquid in a tank. The pressure in the tank is about 800–900 lb/in.^{2} (5500–6200 kilopascals); small leaks are therefore serious, and the tank should be weighed frequently to ensure that the contents are still intact. Because of the high pressure, large carbon dioxide containers are not practical. For small, hand-carried extinguishers, the carbon dioxide tank has a short hose with a horn at the end through which the gas may be directed onto the fire. Large installations involve a battery of carbon dioxide cylinders that are attached to a manifold line.

In certain industrial applications, distribution systems are used for carbon dioxide analogous to the sprinkler systems for water. For example, when a fire starts in electric power generating stations, jets are opened to blow CO_2 into the generator housings. This quickly blankets the critical area with the gas. Similar equipment is provided in large spray-painting operations.

In some industrial plants, provision is made to use flue gases to extinguish fires. The flue gas must be cooled and washed before it is put on the fire. Steam from a boiler is similarly used in certain special instances, but its application is definitely limited. *See* CARBON DIOXIDE.

Dry chemical extinguishers. A dry powder, consisting principally of sodium bicarbonate, is also used as a fire extinguisher. The powder must have the correct particle size and contain materials that prevent it from caking. The action of the powder is threefold. Probably its most important action is to generate carbon dioxide by the reaction $2NaHCO_3 \rightarrow Na_2CO_3 + H_2O + CO_2$. Because this reaction is caused by the heat of the fire, the carbon dioxide is generated precisely where it is needed. Secondly, the powder serves to cool the burning material. Thirdly, it provides shielding to prevent access of air. Dry chemical is useful for small fires, and especially electrical fires. As added advantages, sodium bicarbonate is nontoxic, does little or no secondary damage, and is easily cleaned up after the fire.

Halogenated hydrocarbons. Carbon tetrachloride, CCl_4, has had a long history as a fire-extinguishing agent. As it is customarily used, in small quantities, the principal action is to supply a heavy blanket of vapor over the fire. Because it is 3.5 times as dense as carbon dioxide, it is very effective for this purpose.

In addition, carbon tetrachloride, in common with all the halogenated compounds, has a definite chemical inhibiting effect on combustion. Even vapor percentages in the air as low as 4.5% will sometimes put out a fire. Larger percentages are usually needed; but the halogenated materials are more effective by far in suppressing the inflammability of gasoline vapor, for example, than inert gases such as carbon dioxide and nitrogen.

Other halogenated hydrocarbons that have been used for fire extinguishing besides CCl_4 are chlorobromomethane, CH_2ClBr, and several of the fluorinated hydrocarbons known as freons. An effective agent that has been of interest is Freon 13B-1, CF_3Br.

The principal difficulty, however, with all the halogenated hydrocarbons, and with CCl_4 in particular, is toxicity. The vapors of carbon tetrachloride are themselves quite dangerous. They may render a person unconscious and prevent escape from the fire. More serious still are the gases that are produced from CCl_4 by reactions in the fire. In addition to other noxious fumes, small percentages of phosgene, $COCl_2$, may be generated, and this gas is most deadly. Certain of the freons are much less toxic than CCl_4 and less apt to give dangerous products. *See* HALOGENATED HYDROCARBON.

Other extinguishing methods. Small fires are best treated by choking off the supply of air. In addition to the several methods for doing this already described, others should be mentioned that require no special equipment. For a household fire, especially involving clothing on a person, a blanket or a rug provides an effective means to smother the fire. Small fires around a kitchen stove may be snuffed out with salt or, better still, with bicarbonate of soda. A bucket of sand, strategically located, is also useful against domestic fire hazards.

A very novel technique for putting out a fire burning at the liquid surface in an oil tank or vat is to supply agitation by blowing air in at the bottom of the vessel. (A built-in means must be provided for doing this.) As a result of the agitation, the cold oil from the bottom is brought up to displace the hot burning layer of oil at the top. This method, where applicable, is amazingly effective.

Still other firefighting methods are called for in various special cases. One of these is the explosion of firedamp (methane) in mines. The explosion generates a wind in the mine shaft because the air is pushed ahead of the flame. This accelerates the progress of the explosion and increases the devastation. It has been found that stone dust will stop such fires when it is spread along the mine passages so as to be stirred up readily by the wind. The stone dust probably cools the hot gases to the point where the flame will not propagate.

In the case of forest fires, different techniques, such as back-firing, are used. These depend on wind conditions, moisture, and fuel type. Water suspensions of sodium calcium borate have proved effective in preparing a fire break and even in direct attack on the fire. Borates are well known as fire retardants for paper and fabrics. *See* COMBUSTION; EXPLOSIVE; FIRE DETECTOR; FLAME; FLAMEPROOFING; FOREST FIRE. William E. Gordon

Bibliography. National Fire Protection Association, *Fire Protection Handbook*, 18th ed., 1997; C. V. Walsh and L. Marks, *Firefighting Strategy and Leadership*, 2d ed., 1976.

Fire technology

The application of results of basic research and of engineering principles to the solution of practical fire protection problems, but entailing, in its own right, research into fire phenomena and fire experience to provide an improved level of understanding of the nature of unwanted fire and its threat to life and property. The defense against destructive fire has been largely empirical and improvisational; although this approach has been reasonably effective in the past, it has not kept pace with urbanized and industrialized life. However, an impressive array of scientific and engineering disciplines has been brought to bear on the complex practical problems of fire defense. This article reviews the current state of fire technology, principally as it applies to urban fires.

Fire Fighting and Protection

The fundamental techniques used by fire fighters consist primarily of putting water on a fire. Water serves to cool a burning material down to a point where it does not produce gases that burn. While water is the most practical and inexpensive extinguishing agent, modern technology has provided not only additives to water to render desirable properties such as easy flow or enhanced sticking, but also chemical and physical extinguishants such as fluorocarbons, surfactant film-forming proteins, and foams. *See* FIRE EXTINGUISHER.

Fog or spray nozzles. Water cools more efficiently if it is broken up into tiny droplets which expose more water surface to heat, hence the increasing popularity of the so-called fog or water-spray nozzles. Wetting agents and other surfactant materials are also employed to make water more efficient.

With the introduction of fog and spray nozzles, there is a tendency to use smaller hose lines, principally the $1^1/_2$-in. (38-mm) booster lines, and smaller quantities of water to fight fires in dwellings, apartments, and other smaller occupancies. Fog nozzles have been used successfully on nearly all types of fires, even oil, tar, grease, and electrical fires where water had previously been considered ineffectual or dangerous. It also has been found that water spray or fog absorbs some of the toxic gases of the fire and tends to clear away the smoke. As a result, the sprays are widely used in fighting smoky interior fires.

Fire fighters generally agree that the way to fight a fire is to get inside the building, find the source, and get water directly on it. However, when alarms are delayed and the fire is fully developed before the arrival of the fire fighters, hose streams may be required to reduce the fire to the point where it can be approached. Among various modern developments in water sprinklers, the most noteworthy is the "on-off" system which turns itself off when the fire is under control and comes on again if the fire rekindles.

Ventilation. An important fire-fighting tactic is that of building ventilation, accomplished by the fire fighters' practice of chopping holes in roofs and breaking windows in portions of the building which do not appear to be involved in the fire. This tactic is used to ventilate the building and clear it of heavy smoke and heated gases so that the building can be entered and the fire reached. Ventilation reduces hazards to life and damage due to smoke and heat. Ventilation should be attempted only by experienced fire officers since improper ventilation can spread the fire and increase the damage and can even introduce the risk of explosion. Fire departments may also use electrically driven fans to help clear buildings of smoke.

Fire Prevention and Loss Reduction

The contribution of the practices of fire prevention is potentially much greater than that of the actual fire-fighting activities. Fire-prevention and loss-reduction measures take many forms, including fire-safe building codes, periodic inspection of premises, fire-detection and automatic fire-suppression systems in industrial and public buildings, the substitution of flame-retardant materials for their more flammable counterparts, insurance to offset the financial loss of fire, and the investigation of fires of suspicious origin, serving to deter the fraudulent and illegal use of fire. *See* AUTOMATIC SPRINKLER SYSTEM; FIRE DETECTOR.

The revision of building codes, the establishment of meaningful fire-resistive tests, and many other activities of fire-protection engineering constitute a field in themselves. Similarly, the subject of fire insurance is highly specialized and complex. There is much interest in improving methods for investigating fires of unknown origin. Fire insurance companies employ or retain investigators to determine the cause of a large number of fires of suspicious origin each year. In addition, many fire fighters are trained to recognize the clues pointing to the origin of a fire.

Fire Problems of the Future

Fire involves the chemical reaction between oxygen (air) and a fuel which is raised to its ignition temperature by heat. This definition may be depicted by the fire triangle (**Fig. 1**), which shows the three ingredients required for a fire. The reduction or removal of a sufficient amount of any one of these will stop the fire. This concept formed the rudimentary beginnings of the technology of fire control.

Classification of fires. Fires are divided into four classes according to the type of combustible involved.

Class A. Fires involving solids such as wood, paper, and most plastics are termed class A fires and were almost the only type that used to concern people. Water is the most effective, the cheapest, and the most easily applied extinguishing agent for class A fires.

Class B. These are fires of organic liquids such as gasoline, fuel oil, benzene, and acetone. Such fires became important with the advent of the petroleum industry and other modern chemical processing industries. These fires can be controlled by removing the air with a blanketing agent, such as carbon dioxide or a water-based foam. Water in the form of high-pressure fog is an effective extinguisher, but conventional water streams are unsuitable because they spread the fire.

Class C. A class C fire is any fire in which energized electrical equipment is involved. Because of the hazard of electric shock, any extinguishing agent may be used that does not form electrically conductive paths, including high-pressure water fogs. Aqueous foams are not suitable.

Class D. Fires fueled by sodium, magnesium, titanium, and other reactive metals or their reactive compounds, such as sodium hydride, compose this class and are latecomers to the fire scene. These fires are best controlled by removal of air by a blanket of unreactive powder, such as sodium chloride or graphite. Carbon dioxide, sand, and water cannot be used either to blanket or to cool reactive-metal fires because these substances are a source of oxygen for these fuels.

Future hazards. Parodoxically, some of the remedies devised to alleviate one aspect of the fire problem may be adding worse problems in another area. A case in point is the use of fire-retardant formulations in synthetic and natural polymers. Some of these alter the pyrolysis chemistry, producing more toxic effluents. Increased use of these flame retardants may actually increase the life hazards in a fire.

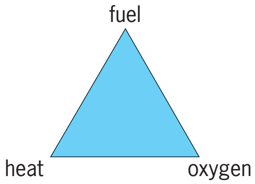


Fig. 1. Triangle of essential ingredients of a fire.

So far, fire has been discussed in terms of the fuel involved without regard to the oxidant because it has been assumed that the oxidant is supplied by air. Fires involve oxidizing agents other than the oxygen in air. Oxidizers such as oxygen, fluorine, and chlorine as liquified gases, and concentrated solutions of hydrogen peroxide find wide industrial use and are transported by highway trucks and railroad cars in growing quantities, entering into the urban scene with increasing regularity. These materials are hazardous because they oxidize, or burn, many common materials of construction, such as wood, paints, plastics, and rubber, and because the temperature required to ignite the combination may be considerably lower than that required for a fire in air alone. Finally, increased use of concentrated oxygen atmospheres and high-pressure, or hyperbaric, atmospheres constitutes a new danger.

Control tactics. The various tactics used in fire prevention and control essentially involve modification of the fire characteristics to alter the time factor to save lives and minimize property damage. Such modifications include preventing fire by reducing possible ignition sources and increasing the ignition time to infinity; early detection of the fire and prompt sounding of the alarm; reducing the fire department arrival time; facilitating quick attack, easy egress, and rescue; decelerating the fire growth; delaying the flashover; and enhancing structural fire endurance. As the National Fire Protection Association (NFPA) points out, fire-safe construction starts on the drafting board, where careful attention can be given to such complexly interrelated factors as the selection of materials, details of architecture, proposed use of the structure, building codes, and legal requirements. All too often, fulfillment of the minimum safety requirements of the codes is no guarantee against a disaster, for what goes into the building changes with time and people.

Fire safety strategies. Of the many specific technological strategies to improve fire safety in an urban structure, the 10 most important are discussed below.

Location and exposure control. This requires placement (zoning) of structures in a community as well as placement of hazardous processes on the periphery of a structure or area. Factors to be taken into account include not only the combustibility of construction materials and the hazards of occupancy, but also the dangers of possible ignition of the building or its contents from exposure to a fire in an adjoining building or other exterior source of ignition.

Occupancy control. This strategy advocates either prohibition or restricted entry of certain materials into a structure with regard to their life hazard and combustibility.

Early detection and alarm systems. The sooner a fire is detected, the easier it is to put out and the less the threat to the occupants.

The five principal types of fire detectors in common use are:

1. *Electronic combustion-gas detector*. This device detects the combustion gases at an early stage of a fire, often long before the smoke would be noticed. Such a detector can usually monitor in excess of 500 ft^2 (50 m^2 or more) of floor area and is preferred in occupancies where life safety is critical.

2. *Radiant-emission detector*. Such a detector responds to infrared or ultraviolet radiation. In some cases, this detector is designed to respond to rapid fluctuations in the radiation, the rate of fluctuations being chosen to correspond approximately to the flicker rate of fires but to exclude the background radiation from other sources. Some versions have a response delay to further screen out background signals. Such a detector can monitor as much as 10,000 ft^2 (1000 m^2) of floor area if its field of view is unobstructed.

3. *Light-scattering (or refracting) detector*. This device appears in various versions, each employing a different principle. One commonly used detector of this type is based on the scattering of light by smoke particles. A new device uses the refraction of a multiply reflected laser beam to detect convection cells produced by a fire. The smoke detectors are relatively slow, and they sample only small volumes of room air, but the laser-beam approach appears to suffer neither disadvantage. *See* LASER.

4. *Fixed-temperature detector*. This detector triggers an alarm upon reaching a preset temperature. Typical versions respond at a temperature level of 158°F (70°C). They are also available in temperatures ranging from as low as 122°F (50°C) to as high as 266°F (130°C). The fixed-temperature detector ordinarily will protect an area of about 100 ft^2 (10 m^2). Relatively sluggish in response, this type of detector is unsuited to life safety and is usually employed in protection of unoccupied property, such as in automatic sprinkler systems for warehouses and the like.

5. *Rate-of-temperature-rise detector*. Such a type responds to a quick rise in room temperature. A typical version responds when the rate of rise exceeds about 0.4°F/s (0.2°C/s); first, however, the temperature must exceed a certain fixed value. A rate-of-rise detector usually monitors about 200 ft^2 (20 m^2) of floor area.

Fire loading control. This strategy could be a part of occupancy control. It involves limiting the maximum amount of combustible material allowed per unit floor area of a structure. Fire loading does not tell much about the ease of ignition and speed of fire growth, but in conjunction with the geometry of the structure, it tells something about the severity of the fully involved fire. The quantity of fuel and its rate of burning determine the duration of a fire.

Regulation of materials. A structure built of steel, concrete, and brick is, of course, safer than one built of wood. When wood is used, application of gypsum or asbestos board improves the safety. The fire properties of building materials fall broadly into two categories: combustibility and fire resistance. Combustibility is further divided into ease of ignition, fire spread, fuel contributed, and smoke developed. The fire resistance of materials and assemblies is expressed in hours of endurance to the exposure from a standard time-temperature furnace.

Compartmentation, subdivision, and ventilation. An effective way of restraining spread of fire through a structure is to suitably seal off each unit of the structure with walls, ceilings, floors, windows, and doors which are impervious to fire. Such a practice is frequently followed in large modern buildings subdivided by fire walls and fire curtains.

In subdivision of buildings, factors must be considered such as the occupant population and the available modes of egress. Pressurized stairwells are under consideration to facilitate safe egress atmospheres. The concept of ventilation is related peripherally to compartmentation and subdivision in that the sealed atmosphere of a modern high-rise building requires, upon detection of a fire, automatic opening of certain vents and closing of others to keep smoke and toxic gases from entering the entire air circulation system.

Structural stability and fire integrity. Since a modern structure is a complex system of many interconnected components, and since a single weak link can defeat the purpose of the entire chain, it is necessary to examine the complex structure for its overall stability and integrity in a potential fire. An otherwise well-designed building and its occupants could, for instance, be placed in jeopardy by improper and inadequate installation of building service equipment.

Accessibility. The need for all-around accessibility of a structure often contradicts the need for security in nonfire situations. Basements, especially, offer some difficulty in accessibility. Exit facilities in accord with the Life Safety Code serve the dual function of providing means of exit from and access to the upper stories of buildings, especially those high-rises that are too tall for fire truck ladders.

Suppression. For most of the problems, advocates in the insurance industry give an all-embracing answer—automatic sprinkler protection. Many commercial and industrial properties are deemed uninsurable unless they are provided with automatically (thermally) actuated water sprinklers. Rigid inspection and maintenance of sprinkler systems are recommended, for insurance company records show that a building with no sprinklers and one with inoperative sprinklers are equally destructible by fire.

Codes, standards, and laws. Codes are, of necessity, generalities, and since fire situations vary over a broad range, literal application of the codes in the design and maintenance of a structure seldom leads to guaranteed safety. Codes and laws are usually based on standards which themselves are deduced from the best available scientific and technological knowledge on fires. Unfortunately, the little available scientific and technological knowledge becomes obscured when filtered through standards to codes. Codes consequently vary widely from place to place in the United States and Canada, but there is a strong movement in the United States to bring uniformity. True uniformity is possible only when enough factual data are developed from science and technology.

The existing codes are in three categories. Life safety codes primarily concern exits—their adequate number and size and unrestricted approach. Building codes provide rules for safe construction. Most municipalities have their own codes, which are based largely on five model codes: National Building Code (NBC), Uniform Building Code (UBC), Southern Standard Building Code (SSBC), Basic Building Code (BBC), and National Building Code of Canada (NBCC). Finally, fire prevention codes exert an important influence on the occupancy and operation of a building by further refining the building codes.

Chemically active fire-suppression agents. A general approach to fire control has been developed involving use of flame inhibitors. Unlike older fire-extinguishing materials such as water and carbon dioxide, these agents operate indirectly in that they interfere with those reactions within a flame that lead to sustained release of heat. As a result, temperature of the system falls below ignition temperature.

Chemical flame inhibitors offer one of the greatest hopes for better control of fires of all sizes. These agents have achieved such importance that they have, in effect, added a new dimension to the fire triangle, making it a tetrahedron, with one additional mechanism for interfering with the combustion process (**Fig. 2**).

Chemical extinguishers are of two types: liquid (or liquefied gas) and dry powder.

Liquid chemical extinguishers. The most effective liquids are the halogenated hydrocarbons such as chlorobromomethane (CB) and bromotrifluoromethane (better know as Halon 1301) which are colorless, odorless, and electrically nonconductive. *See* HALOGENATED HYDROCARBON.

Automatic Halon 1301 extinguishing systems are ideally suited to situations in which damage of equipment or furnishings by water or dry chemicals would cause irreparable harm. It is used in automatic fire protection systems by insurance companies, banks, universities, and other organizations with computer facilities and valuable, irreplaceable records on data-processing cards or tapes to protect equipment, subfloor electrical wiring, and record libraries.

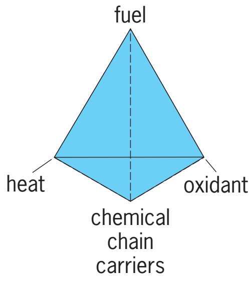


Fig. 2. Chemical fire suppressants add another dimension to the fire triangle, forming a fire tetrahedron.

Tests with Halon 1301 have been successful in hyperbaric (pressurized) atmospheres as well. The atmospheres in hyperbaric chambers are oxygen-rich if not pure oxygen. Materials that are fire-resistant in normal air burn enthusiastically in such oxygen-rich environments, but can be extinguished by Halon 1301 if they can be detected in time.

Dry-powder chemical extinguishers. Ammonium dihydrogen phosphate is the most useful of this type and is rated for class A, B, and C fires. Other dry-powder inhibitors are salts of alkali metals (which include lithium, sodium, potassium, rubidium, and cesium). For example, the dry powders based on sodium bicarbonate and the more effective potassium bicarbonate are rated as extinguishers for class B and C fires. The rubidium and cesium salts have not been studied in detail, and lithium salts have no appreciable activity. Dry chemicals can be effective even when they are applied as a fog in water solutions.

Water. Notwithstanding the success of chemical inhibitors, water is still the prime suppressant because of its great cooling power, general availability and low cost. However, application in a solid stream is regarded as wasteful of water and damaging to the components and contents of structures. Water spray and fog have been highly effective, as mentioned above, even in quenching gasoline fires, for which water is normally unsuitable. Moreover, spray and fog minimize water damage.

Foams. Foams are also widely used. Protein-type, low-expansion foams, particularly useful in quenching burning volatile petroleum products, are used in crash-rescue operations. High-expansion foams are available for fire suppression in enclosed areas. Some foams of this type are generated at a rate of 15,000 ft^3/min (424.8 m^3/min). They can contain sufficient air to allow a human to breathe inside of them.

Light water. A film-forming solution of a specific fluorocarbon surfactant in water, known as light water, was developed by the U.S. Navy for use with dry chemicals to fight aircraft crash fires. It may be used either as a liquid or as a low-expansion foam to interfere with the release of flammable vapors from the burning fuel. Light water is also useful in extinguishing petroleum storage tank fires and may find application to urban fires once the cost is no longer prohibitive. Staney B. Martin; A. Murty Kanury

Bibliography. R. E. Colburn, *Fire Protection and Suppression*, 1975; R. L. P. Custer and R. G. Bright, *Fire Detection: The State of the Art*, NBS Tech. Note 839, 1974; W. M. Haessler, *The Extinguishment of Fire*, NFPA, rev. ed., 1974; H. E. Hickey, *Hydraulics for Fire Protection*, 1980; R. F. Mendes, *Fighting High-Rise Building Fires*, National Fire Protection Association, 1975; National Commission on Fire Prevention and Control, *America Burning*, 1973; National Fire Protection Association, *Fire Protection by Halons*, collection of articles from *Fire Journal* and *Fire Technology*, NFPA SPP-26, 1975; National Fire Protection Association, *Inspection, Testing and Maintenance of Sprinkler Systems*, 1987.

Fire-tube boiler

A steam boiler in which hot gaseous products of combustion pass through tubes surrounded by boiler water. The water and steam in fire-tube boilers are contained within a large-diameter drum or shell, and such units often are referred to as shell-type boilers. Heat from the products of combustion is transferred to the boiler water by tubes or flues of relatively small diameter (about 3–4 in. or 7.5–10 cm) through which the hot gases flow. The tubes are connected to tube sheets at each end of the cylindrical shell and serve as structural reinforcements to support the flat tube sheets against the force of the internal water and steam pressure. Braces or tension rods also are used in those areas of the tube sheets not penetrated by the tubes.

Fire-tube boilers may be designed for vertical, inclined, or horizontal positions. One of the most generally used types is the horizontal-return-tube (HRT) boiler (**Fig. 1**). In the HRT boiler, part of the heat from the combustion gases is transferred directly to the lower portion of the shell. The gases then make a return pass through the horizontal tubes or flues before being passed into the stack.

In the Scotch marine boiler, one or more large flues (approximately 18–24 in. or 45–60 cm in diameter) are used for furnaces or combustion chambers within the shell, and the hot gases are returned through groups of small-diameter tubes or flues. The flues that form the combustion chamber are corrugated to prevent collapse when subjected to the water and steam pressure. These boilers may be oil-fired, or solid fuel can be burned on grates set within the furnace chambers. Scotch marine boilers have, with few exceptions, been superseded by water-tube marine boilers.

Gas-tube boilers are sometimes used for the absorption of waste heat from process gases or the exhaust from internal combustion engines, particularly when their installation provides a simple and economical means for the recovery of low-grade heat. The boiler shell may be augmented by an external steam-separating drum and downcomer and riser connections to provide for proper circulation (**Fig. 2**).

Shell-type boilers are restricted in size to 14 ft (4.3 m) in diameter and an operating pressure of 300 lb/in.^{2} (2.1 megapascals) by the stresses in the large-diameter shells and the necessity to design the flat-tube sheets for practicable thicknesses. They are best suited for low-pressure heating service and for portable use because the integrated structure can be transported easily and the generous reserve in water capacity requires minimum attendance. However, there is the risk of catastrophic damage if overheating and rupture of the shell or tube sheet occur as a result of low water levels or the formation of an insulating layer of internal scale produced by water impurities. *See* BOILER; STEAM-GENERATING UNIT.

George W. Kessler

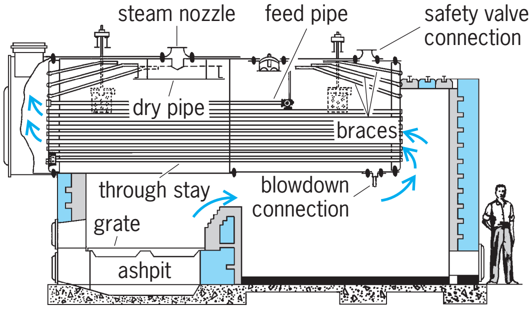


Fig. 1. Horizontal-return-tube boiler.

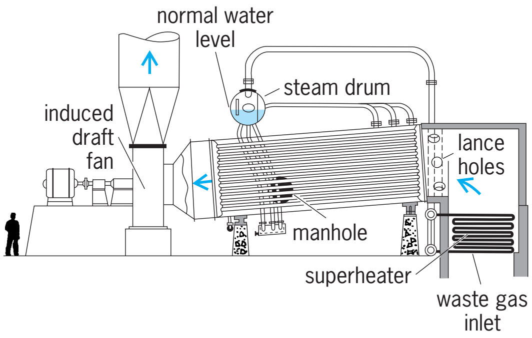


Fig. 2. Fire-tube waste-heat boiler.

Fired heater

A direct-fired heater used in processing petroleum. The basic feature of these heaters is that the fluid (liquid, vapor, or a mixture of both phases) is heated in carbon- or alloy-steel tubes, which in turn receive their heat from hot products of combustion. The fuel may be liquid, such as heavy fuel oil, or gas, such as natural gas or cracked refinery gas. Burners of the steam, gas, or mechanically atomizing type may be used for liquid fuels, and burners of the simpler ring-type, spider, or center-nozzle type for gas fuels. Combination burners designed to handle both fuels are in wide use.

The fired heater is also known as a tube-still heater. Use of "still" in the name comes from the fact that the first continuous heaters were those for heating the topping still, which separated the various fractions of the virgin crude by distillation at atmospheric pressure. Later, to permit distillation of higher-boiling fractions, the distillation was carried out under subatmospheric pressure, whence the name vacuum-pipe still. With the advent of thermal cracking, the tubular heater served the dual purpose of heating the oil to be cracked to the reaction temperature and of maintaining it under the proper temperature, pressure, and time conditions while providing the required endothermic heat of reaction. Today it is the general-purpose tool for supplying heat at the higher temperature levels, that is, over 650–700°F (343–371°C), wherever required in the refinery for distillation, for preheating, for reactions such as pyrolysis, or for reboiling.

Petroleum heaters may heat a single fluid stream, such as the crude charge to a unit. In this case, the

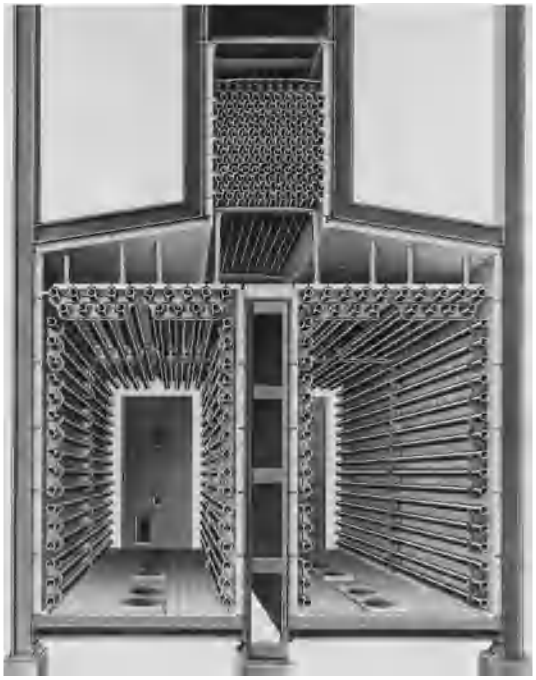


Fig. 1. Double radiant box furnace. (*Lummus Co.*)

tubes are connected in series flow by suitable return bends, which may be provided with cleanout openings. Two or more parallel fluid streams may be heated within the same box in order to keep the friction drop through the heater down to desirable limits. Sometimes streams for different services are heated in the same furnace. However, flexibility today usually dictates that the streams be fired and controlled separately. In certain cases, for example, in steam-methane reformers, where large volumes of gas are passed through catalyst-filled tubes and pressure drop must be minimized and temperatures carefully controlled, all furnace tubes are manifolded in parallel, with all tubes given essentially the same treatment.

Most modern heaters are of one of two types, the box type, usually with horizontal tubes (**Fig. 1**), or the cylindrical type with vertical tubes arranged in a circle (**Fig. 2**). Within the radiant section, the combustion zone in which the burners are located, heat is given up to the tubes from four sources: radiation from the flame, radiation from the hot molecules (carbon dioxide and water vapor), radiation from the gases within the chamber. The convection section is basically a recuperative zone in which the flue gases can be reduced in temperature while heating the coldest fluid to obtain economic heat recovery. Higher overall efficiency may be obtained with air preheaters. In this latter zone, heat is transferred to the air used in combustion by radiation from the hot gases passing through the staggered bank of tubes, by convection from such gases and by radiation from the hot refractory confining the section. The first two rows to receive the hot flue gases from the radiant section often are called the shield or shock tubes, because, as a result of their position, they also receive heat either directly from the combustion box, as do the radiant tubes, or indirectly from large areas of hot radiating refractory.

The design of fired heaters is based as much on experience as on engineering science. The designer must consider not only structural arrangement and heat-transfer rates, but also the effects of fly ash on refractories and metals, strains from heating and cooling, corrosion and erosion, inlet and outlet pressures, methods of operating, and methods of control.

The calculation of pressure drop on the fluid side of the tubes is often complicated because the conditions are nonisothermal, and the relative volumes and compositions of the liquid and gas phases may be changing rapidly.

Tube stills have been designed for operating temperatures up to 1900°F (1037°C), for pressures as high as 4500 lb/in.^{2} (30 megapascals), and as low as 60 mmHg (8 kilopascals). Radiant-section heat-transfer flux may be as high as 30,000–40,000 Btu/(h)(ft^2) [95,000–126,000 W/m^2],

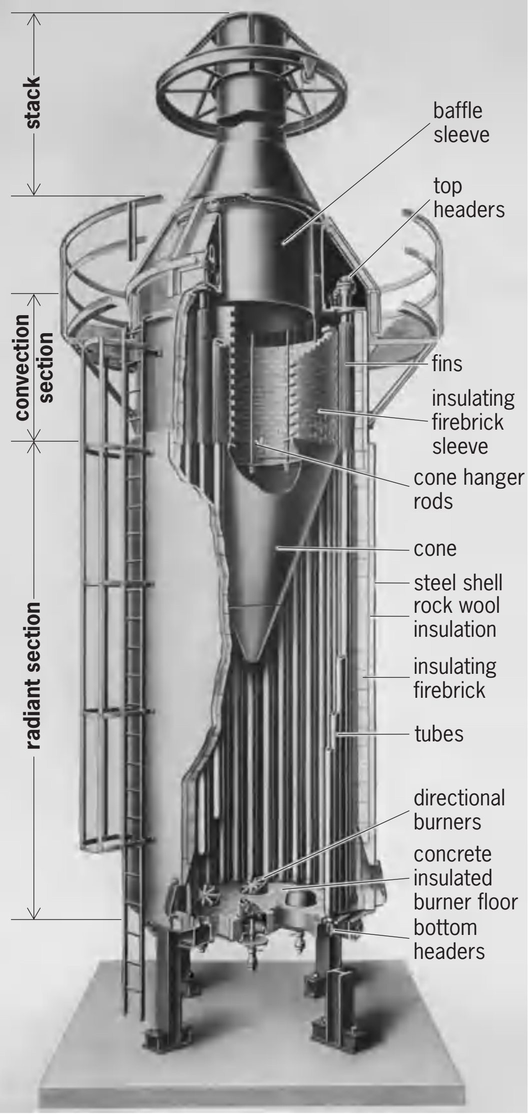


Fig. 2. Cylindrical-type furnace. (*Petro-Chem Development Co.*)

though the more common figure is 12,000–14,000 Btu/(h)(ft^2) [38,000–44,000 W/m^2], for the cleaner, lighter oils and 8000–10,000 Btu/(h)(ft^2) [25,000–32,000 W/m^2] for the heavier, dirtier ones. The severe conditions often encountered in furnace tubes were a prime incentive in the earlier development of strong, high-temperature alloys. *See* FURNACE CONSTRUCTION; HEAT TRANSFER; PETROLEUM PROCESSING AND REFINING. Walter E. Lobo

Fischer-Tropsch process

The synthesis of hydrocarbons and, to a lesser extent, of aliphatic oxygenated compounds by the catalytic hydrogenation of carbon monoxide. The synthesis was discovered in 1923 by F. Fischer and H. Tropsch at the Kaiser Wilhelm Institute for Coal Research in Mülheim, Germany. The reaction is highly exothermic, and the reactor must be designed for adequate heat removal to control the temperature and avoid catalyst deterioration and carbon formation. The sulfur content of the synthesis gas must be extremely low to avoid poisoning the catalyst. The first commercial plant was built in Germany in 1935 with cobalt catalyst, and at the start of World War II there were six plants in Germany producing more than 4,000,000 bbl (6,000,000 m^3) per year of primary products. Iron catalysts later replaced the cobalt.

Following World War II, considerable research was conducted in the United States on the iron catalysts. One commercial plant was erected at Brownsville, Texas, in 1948, which used a fluidized bed of mill scale promoted by potash. Because synthetic oil was not competitive with petroleum, the plant was shut down within a few years. Fischer-Tropsch plants constructed in South Africa (SASOL) have produced gasoline, waxes, and oxygenated aliphatics. The SASOL plants gasify inexpensive coal in Lurgi generators at elevated pressure. After purification, the gas is sent to entrained and fixed-bed synthesis reactors containing iron catalysts. *See* COAL GASIFICATION; COAL LIQUEFACTION.

Methanation. In the years since 1960, interest has grown in the United States in catalytic methanation to produce high-Btu gas from coal. Synthesis gas containing three volumes of hydrogen to one volume of carbon monoxide is reacted principally to methane and water, as shown in reactions (1). Nickel catalysts,

$$\begin{aligned}(2n+1)H_2 + nCO &\xrightarrow{\text{Co catalysts}} C_nH_{2n+2} + nH_2O\\ (n+1)H_2 + 2nCO &\xrightarrow{\text{Fe catalysts}} C_nH_{2n+2} + nCO_2\end{aligned} \qquad (1)$$

first discovered by P. Sabatier and J. B. Senderens in 1902, are important catalysts for methanation. Active precipitated catalysts and Raney nickel catalysts were developed, and the Bureau of Mines developed a technique for applying a thin layer of Raney nickel on plates and tubes by flame-spraying the powder. This led to development of efficient gas recycle and tube-wall reactors.

Reactions. Typical Fischer-Tropsch reactions for the synthesis of paraffins, olefins, and alcohols are reactions (1), (2), and (3), respectively.

$$\begin{aligned}2nH_2 + nCO &\xrightarrow{\text{Co catalysts}} C_nH_{2n} + nH_2O\\ nH_2 + 2nCO &\xrightarrow{\text{Fe catalysts}} C_nH_{2n} + nCO_2\end{aligned} \qquad (2)$$

$$\begin{aligned}&2nH_2 + nCO \xrightarrow{\text{Co catalysts}}\\ &\qquad C_nH_{2n+1}OH + (n-1)H_2O\\ &(n+1)H_2 + (2n-1)CO \xrightarrow{\text{Fe catalysts}}\\ &\qquad C_nH_{2n+1}OH + (n-1)CO_2\end{aligned} \qquad (3)$$

The primary reaction on both the cobalt and the iron catalysts yields steam, which reacts further on iron catalysts with carbon monoxide to give hydrogen and carbon dioxide. On cobalt catalysts, at synthesis temperatures (about 392°F or 200°C), the reaction $H_2O + CO = CO_2 + H_2$ is much slower than on iron catalysts at synthesis temperatures (482–617°F, or 250–320°C). All these synthesis reactions are exothermic, yielding 37–51 kilocalories/mol (160–210 kilojoules/mol) of carbon in the products or 4700–6100 Btu/lb of product (11,000–14,000 kilojoules/ kg).

The hydrocarbons formed in the presence of iron catalysts contain more olefins than those formed in cobalt catalyst systems. The products from both catalysts are largely straight-chain aliphatics; branching is about 10% for C_4, 19% for C_5, 21% for C_6, and 34% for C_7. Aromatics appear in small amounts in the C_7 and in larger amounts in the higher boiling fractions. Operating conditions and special catalysts required, such as nitrides and carbonitrides of iron, have been ascertained for the production of higher proportions of alcohols. Joseph H. Field

Bibliography. R. B. Anderson, *The Fischer-Tropsch Synthesis*, 1984; *Kirk-Othmer Encyclopedia of Chemical Technology*, 3d ed., vol. 11, 1998; H. H. Lowry (ed.), *Chemistry of Coal Utilization*, vols. 1 and 2, 1945, suppl. vol., 1963.

Fisher

A medium-sized, weasellike, carnivorous mammal classified in the family Mustelidae along with weasels, mink, martens, wolverines, otters, and badgers. Fishers (*Martes pennanti*) occur only in North America. They are found in forested areas of northern North America and southward in the Sierra Nevada Mountains to central California; and southward in the Appalachian Mountains to West Virginia. Their original range, which has been significantly reduced, particularly in the United States, nearly coincided with the combined distributions of northern hemlock, northern hardwood, western mountain, and boreal forests. The fisher is known sometimes as the Pekan, Pennant's marten, black cat, or American sable, although fisher appears to be the most popular name. Oddly enough, the animal does not fish.

Fisher, *Martes pennanti.* (*From Living with Fisher in Massachusetts,* © *Mass Wildlife*, *www.mass.gov./dfwele/dfw/dfw_fisher.htm*)

General morphology. The fisher has a long body with short legs, small rounded ears, and a long, bushy tail. The head is wedge-shaped and tapers to a narrow black nose. The body, which generally resembles that of a stocky weasel, is set low to the ground. The pelage (coat), which is dark brown to nearly black, is long except on the face. It differs between the sexes and changes with the seasons. The pelage of females is darker and silkier than that of males, whose fur is coarser and more grizzled. During winter, the tail, rump, and legs are glossy black. White-tipped guard hairs give the face, neck, and shoulders a frosted appearance. Molting occurs in late summer and autumn. The feet are large with each foot consisting of five toes with sharp, curved, unsheathed claws. The large feet, which are heavily furred in winter, presumably assist in walking on snow. Locomotion is digitigrade (that is, they walk on the digits). They can swim if necessary. Both sexes possess a pair of lateral anal sacs that contain a brown, musky-smelling, viscous liquid. The dental formula is I 3/3, C 1/1, Pm 4/4, M 1/2 × 2 for a total of 38 teeth. Males are larger, heavier, and more muscular than females. Adult head and body length is 490–639 mm (19 − 25 in.), and tail length is 253–425 mm (10 − 16 in.). Males weigh 2.6–5.5 kg (5.6 − 12 lb), and females weigh 1.3 − 3.2 kg (2.8 − 7 lb). *See* DENTITION.

Habitat. The fisher inhabits large, heavily wooded areas consisting of spruce, fir, or mixed hardwood trees. It also inhabits timbered northern bogs and swamps. It avoids open places with no overhead cover such as fields, roads, burns, and open bogs. Temporary dens for cover and protection may be located under logs, brush piles, or tree roots, in hollow trees or ground burrows, or under the snow. Temporary dens are seldom used for more than 2 or 3 days. Nesting dens occur most often high in hollow trees.

Fishers are shy and elusive animals that are rarely seen, even in areas where they are abundant. They are most active at night, but daytime activity is not uncommon. They are active year-round. They climb well and, although they spend a considerable amount of time foraging and resting in trees, their activity is predominantly terrestrial. They descend trees head first like a squirrel with their hindlimbs rotated outward and caudally. They are largely solitary, except during the breeding season, with members of each sex defending territories only against members of their own sex. Male territories may range from 20 to 80 km^2 (7.7 to 31 mi^2), whereas female territories range from 4 to 35 km^2 (1.5 to 13.5 mi^2). Territory size depends on habitat quality and prey availability.

The fisher is a generalized, opportunistic predator. Mice, squirrels, hares, rabbits, porcupines, birds, and even carrion are the predominant foods, although birds and their eggs, insects, amphibians, reptiles, and various fruits and nuts are consumed when available. The fisher is one of the few predators to successfully prey on porcupines. When a porcupine is found on the ground, the fisher attacks its face, which is not protected by quills. After 30 min or more, it may have bitten the porcupine's face several times, weakening the porcupine sufficiently to turn it over and begin eating from the ventral surface where there are few quills.

Reproduction and development. Although the fisher breeds in March or April, implantation of the zygote is delayed for about 11 months until the following February or March, after which an approximately 30-day active gestation ensues. Thus, total gestation ranges from 327 to 358 days. The average litter of three young is born from February to May, with the female breeding again within 10 days. Thus, except for a few days following parturition, adult females are pregnant almost all the time. Maternal dens are located in a tree cavity high in a large tree and are used for the first 8–10 weeks. Once the kits become mobile, they are moved from the maternal den to one on or below the ground. It is believed that the maternal den functions to protect the helpless young from aggressive male fishers and ground predators. Newborn young are helpless, blind, and sparsely furred (that is, they are altricial). The eyes open at about 53 days, and they venture outside the den for the first time at about 10–12 weeks of age. Weaning occurs at about 15 weeks. The young travel with the female until late summer or early fall, at which time they disperse to begin their solitary lives. Females are sexually mature when a year old, produce their first litters when they are 2 years old, and breed annually thereafter. Fishers have lived 10 years in the wild and over 10 years in captivity. Humans, by hunting and trapping, are the only predator of adult fishers, which are highly valued for their fur. Kits may be preyed upon by hawks, owls, foxes, bobcats, lynxes, and black bears. Overharvesting and habitat destruction have resulted in the fisher's extirpation from much of the southern portion of its original range. Restocking has been attempted in a number of areas with varying amounts of success. *See* CARNIVORA; DENTITION. Donald W. Linzey

Bibliography. D. W. Linzey, *The Mammals of Virginia*, McDonald & Woodward, Blacksburg, Virginia, 1998; R. A. Powell, *The Fisher: Life History, Ecology, and Behavior*, 2d ed., University of Minnesota Press, Minneapolis, 1982; R. A. Powell, *Martes pennanti*, *Mam. Species*, no. 156, American Society of

Mammalogists, 1981; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Fisheries ecology

The study of ecological processes that affect exploited aquatic organisms, in both marine and freshwater environments. Because this field is primarily motivated by an attempt to harvest populations, special attention has been given to understanding the regulation of aquatic populations by nature and humans. The foundations of fisheries ecology lie in population and community ecology. *See* ECOLOGY.

A growing emphasis on the conservation of aquatic resources and on the development of aquatic culture in captivity is leading to a diversification of the issues examined by fisheries ecologists. Ideas, concepts, and methods from physiology, genetics, molecular biology, and epidemiology are increasingly relevant in fisheries ecology. Also, because of the multiplicity of issues associated with harvesting biological resources, fisheries ecology must go beyond biology and ecology into sociology and economics.

Harvesting. The problem of regulating the exploitation of aquatic organisms in order to ensure sustained harvest lies at the core of fisheries ecology. Commercial fishing in the twentieth century showed that unregulated or inadequately regulated harvest frequently results in overexploitation and collapse of fishery systems. During the two decades following World War II, the harvest of marine resources increased threefold. The increase continued as new resources were found and technology magnified the efficiency of commercial fishery. Examples of overexploitation abound, including the collapse of the Peruvian anchoveta in the early 1970s and the collapse of Atlantic cod stocks in the maritime provinces of Canada in the 1990s. Most experts agree that the harvest of marine resources has peaked and increased yields are likely to come only from fine-tuning of regulations on stocks that are fully exploited. Consequently, regulation is becoming more critical if sustained harvests from world fisheries are to be ensured.

The typical unit at which management efforts are directed is the exploited population, customarily termed a stock in fisheries science. A central aspect of assessing fisheries resources is to identify these stocks and determine how isolated they are from other stocks of the same species. For example, there are 20 major recognized stocks of cod in the North Atlantic Ocean, each on the whole isolated from every other and distinct with regard to several biological characteristics that determine the potential for harvest. The three main population processes that govern the size and productivity of given stock are somatic growth, mortality, and recruitment. (Recruitment is a term that was coined in fisheries science to designate the incorporation of new individuals into the population through birth.) A typical assessment of an exploited stock includes the study of these three processes, as well as some protocol to estimate abundance. From this information, and aided by mathematical models and statistical tools, fisheries biologists produce recommendations on how many individuals to harvest, of what size or age, when, and where.

Most of the natural variability in numbers experienced by exploited marine fish populations comes from interannual variability in egg production and survival at the juvenile stages, the two processes that determine recruitment. The study of environmental effects on recruitment is the subject of the field known as recruitment fisheries oceanography. Through the study of the recruitment process, fisheries oceanographers attempt to understand those factors that control abundance and to develop techniques that can help predict future population sizes. Another purpose of fisheries oceanography is the development of tools and procedures to detect, monitor, and analyze changes in biological oceanic processes as a consequence of large-scale environmental changes.

Whereas many fishery studies are performed at the single-species level, exploited populations interact with other species as prey or predators and compete for food or space. For example, krill in the Antarctic oceans is the main food source for numerous species of fish, seabirds, and marine mammals. Its exploitation can thus have far-reaching effects at the community level. Many fisheries, such as those exploiting groundfish in the North Pacific and North Atlantic, have as a target an assemblage of species, each of which may have a different productivity. A key task in these fisheries is to administer the harvest of the conglomerate of species by relatively nonselective fleets.

The seemingly foolish tendency to manage most fisheries from around the world on the basis of single-species approaches is due more to the technical difficulties related to addressing the complexities of species interactions than to a lack of awareness about their significance. In recent years, greater attention has been given to the development of fisheries assessment methods that consider the multispecies nature of exploited systems.

Unlike terrestrial populations, aquatic populations cannot be assessed by direct counts. Most aquatic resources are surveyed indirectly through expensive sampling protocols or from catch data from fisheries. Knowledge about the state of these populations and about their response to exploitation is hence tainted by significant errors. Moreover, a fishery involves not only the biological system but also a social and economic system, with its own rules and behavior. Economic, political, or social realities often overruled wise plans for biological management, leading to prominent failures in fisheries management. Substantial efforts in modern fisheries studies are devoted to finding management strategies and protocols that take into account the uncertainties and complexities associated with exploited populations. *See* ADAPTIVE MANAGEMENT.

Aquaculture. The total world supply of fish has increased steadily in recent years. Much of this positive trend is attributable to aquaculture, which contributed 18.5% of the total world production of finfish and shellfish in 1995 (or 112.9 million metric tons). Aquaculture is the farming of aquatic organisms in fresh, salt, and brackish waters. It is an activity close to agriculture and animal husbandry in its fundamental objective: the intensive production of organisms in controlled and stable environments. That objective is reached by providing adequate quality and quantity of food, shelter from predators, and protection against disease. Fishes, crustaceans, mollusks, and algae are the favorite cultured aquatic species.

Aquaculture activities can be classified by the extent of human intervention in the process. One extreme of that spectrum is supplementation or stocking, by which large numbers of artificially produced individuals are added to natural systems to increase natural production or to mitigate the effects of harvest and other human-induced impacts. This is a widespread practice with salmonids in fresh water and has been the technique used to introduce them around the world, such as in New Zealand, Australia, and South America.

Ranching consists of producing juvenile stages of aquatic organisms in captivity and releasing them into the wild. They are harvested at an older age after reaching commercial sizes as free-ranging organisms. Salmon ranching is a widespread practice in Europe and North America.

Intensive aquaculture is developed through an end-to-end intervention and accounts for most of the world aquaculture production. Organisms are reared from egg to adult in controlled environments. Prominent examples are carp production in Asia and salmon production in marine net pens in Norway and Chile.

While ecological studies to guide harvesting plans emphasize population and community ecology, support for aquaculture activities comes more from those areas of biology and ecology related to physiology, nutrition, genetics and epidemiology. *See* AQUACULTURE.

Conservation. The early history of modern fishing was characterized by a rather utilitarian urge to ensure an increasing supply of aquatic organisms at all costs. In the last two decades, that view has yielded to a more realistic perspective, increasingly heedful to the natural limits of aquatic resources and to many environmental aspects of harvested systems. These include habitat degradation, overfishing, incidental mortality of nontarget species, and the indirect effects of species removal at the ecosystem level.

Habitat degradation is a major threat for aquatic species, especially in fresh-water and coastal systems where human activities are most intensive. Dramatic examples of habitat alterations are found in rivers fragmented by dams. For instance, the construction of hydroelectric projects is one of the prime factors affecting hundreds of salmon stocks at risk in the Pacific Northwest of North America by impeding migration, by flooding prime fresh-water habitat, and by killing fish passing through turbines. Deforestation and erosion of river banks also have significant effects on the survival and reproduction of fresh-water fish, as does the withdrawal of water from rivers for agriculture and pollution by urban and industrial waste. Some fisheries operating on the continental shelf, such as commercial bottom trawling and dredging, produce a profound physical disturbance on sea-floor communities.

The way in which fisheries are conducted or regulated often has important implications for the conservation of species other than those targeted. Incidental capture of accompanying nontarget species (known as "bycatch") is a significant problem in many fisheries, such as in most shrimp fisheries. In the Gulf of Mexico, the ratio of fish discards to shrimp caught ranges from 3:1 to 20:1. In multispecies fisheries, the imposition of catch limits for some species can promote the discarding of otherwise perfectly good fish. In the United States west coast groundfish fishery, discards in excess of 80% of the catch were observed.

Another serious environmental problem associated with some fishing operations is the incidental take of seabirds and marine mammals. Albatrosses and other seabirds in the North Pacific and in the South Atlantic oceans are attracted to baited hooks in longline fisheries for tuna, swordfish and squid as the gear is deployed. When caught on a hook, the birds drown. The impact of incidental mortality on some of these species is yet unknown, but available data indicate that it might be significant. In other cases, the effects of given techniques were unequivocal and led to sweeping regulations, such as the United States ban on driftnet fishing established in 1992. In a few cases, the problem of incidental catch was corrected by relatively simple technological adjustments in the fishing gear or in the fishing operation. The incidental catch of loggerhead turtles in the Gulf of Mexico shrimp fishery, for example, was significantly curtailed by the development of a device that lets turtles escape from fishing nets while retaining shrimp. Similarly, the incidental mortality of dolphins in tuna purse seiners operating in the Pacific Ocean was greatly reduced by a combination of gear modification and changes in fishing protocols.

The indirect effects of removing key species on aquatic communities are extremely difficult to study and detect. Nonetheless, commercial fisheries are suspected to be major degrading factors in some aquatic ecosystems. For example, the decline experienced by populations of marine mammals and seabirds in the Gulf of Alaska and the Bering Sea coincided with the development of major fishing operations in the region. The limitation of forage fish resulting from removals by fisheries has been proposed as a main cause for this decline.

Aquaculture activities can affect the environment and neighboring communities in different ways. Fish waste and excess food from net pens can significantly affect habitat and bottom communities of lakes and coastal areas. Another form of pollution from aquaculture is through the supply of antibiotics to hatchery fish in net pens, which can kill naturally occurring bacteria in bottom communities and have a cascading effect on the rest of the community. In some other cases, the infrastructure required for aquaculture has a substantial impact on the environment. Such is the case of shrimp aquaculture in Latin America, where the construction of rearing ponds has led to the loss of thousands of acres of coastal mangroves.

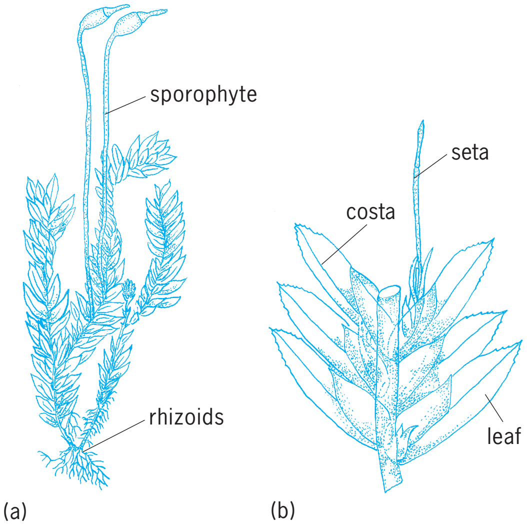
Another undesirable effect of aquaculture is the crossing of escaped hatchery with wild organisms, which can diminish the ability of future hybrid generations to adapt to local wild conditions. Such a situation is found in Norway, the leading producer of cultured salmon in the world. Small escapes from this enormous hatchery production into the wild can dwarf the relatively small wild populations. Not only are the ecological impacts of hatchery fish of concern, but also the danger of introducing diseases that could be deadly for wild fish.

A major problem associated with the development of aquaculture is that it promotes a major trade in exotic species around the world. Species for which culture techniques have been mastered in one corner of the world are being introduced to other regions with the promise to jump-start aquaculture. Often, once they are introduced for aquaculture purposes, individuals of these alien species escape from the confines of culture facilities and establish populations in the wild, with unexpected effects on native communities. A contemporary example is provided by the Atlantic salmon introduced on the Pacific coast of North America, where it has become a favorite species for net pen culture. Atlantic salmon escaping from net pens are regularly found in rivers of the Pacific Northwest, and evidence of successful spawning has been recently gathered. *See* MARINE CONSERVATION; MARINE FISHERIES. Miguel Pascual; Patricia Dell'Arciprete

Bibliography. L. R. Brown et al., *State of the World 1995, Worldwatch Institute*, Norton, New York, 1995; FAO Inland Water Resources and Aquaculture Service, Fishery Resources Division, Review of the State of the World of Aquaculture, *FAO Fisheries Circ.*, no. 886 FIRI/C886 (Rev. 1), Rome, 1997; FAO Marine Resources Service, Fishery Resources Division, Review of the State of the World Fishery Resources: Marine Fisheries, *FAO Fisheries Circ.*, no. 920 FIRM/C920 (En), Rome, 1997; R. Hilborn and C. Walters, *Quantitative Fisheries Stock Assessment*, Chapman & Hall, London, 1992.

Fissidentales

An order of the true mosses (subclass Bryidae) consisting of a single family, the Fissidentaceae, and about 800 species in the genus *Fissidens* plus a few others distributed in four or five segregate genera. *Fissidens* is the largest genus of mosses. The Fissidentales are unique in leaf form and arrangement. The plants, erect and simple or forked, grow (except at the earliest stages) from a two-sided apical cell. The leaves, arranged edgewise to the stem in two rows, have two sheathing blades which clasp the stem and the base of the next leaf above (see **illus.**) Beyond the sheathing blade and toward the stem is a ventral blade; on the side away from the stem is a dorsal blade extending to the leaf base or nearly so. The costa is single and well developed or rarely lacking. The cells are isodiametric and often papillose, though leaf margins are often bordered by linear cells. The archegonia are usually terminal. The capsules are generally exserted, erect, and symmetric, with a peristome that consists of 16 teeth that are characteristically forked and spirally ridged at the tips. The calyptra is cucullate and the chromosome numbers are 5, 6, 8, and 10. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA. Howard Crum

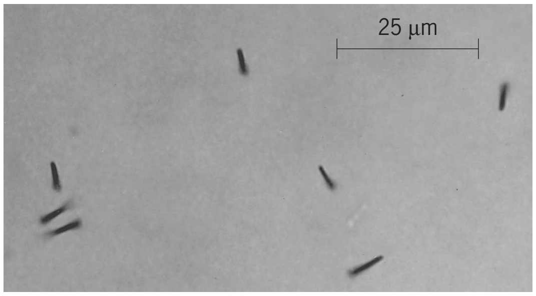


***Fissidens adiantoides*. (*a*) Entire plant. (*b*) Enlarged portion. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

Fission track dating

A method of dating geological and archeological specimens by counting the radiation-damage tracks produced by spontaneous fission of uranium impurities in minerals and glasses. During fission two fragments of the uranium nucleus fly apart with high energy, traveling a total distance of about 25 micrometers (0.001 in.) and creating a single, narrow but continuous, submicroscopic trail of altered material, where atoms have been ejected from their normal positions. Such a trail, or track, can be revealed by using a chemical reagent to dissolve the altered material, and the trail can then be seen in an ordinary microscope. The holes produced in this way can be enlarged by continued chemical attack until they are visible to the unaided eye. *See* URANIUM.

Theory. Track dating is possible because most natural materials contain some uranium in trace amounts and because the most abundant isotope of uranium, ^{238}U, fissions spontaneously. Over the lifetime of



Etched natural fission tracks in muscovite mica crystal. The sample was etched in concentrated hydrofluoric acid.

a rock substantial numbers of fissions occur; their tracks are stored and thus leave a record of the time elapsed since track preservation began. The **illustration** shows etched fission tracks in a mica crystal. The number of tracks produced in a given volume of material depends on the uranium content as well as the age, so that it is necessary to measure the uranium content before an age can be determined. This measurement is most conveniently made by exposing the sample to a known dose of thermal (slow) neutrons, which induce new fissions whose number then gives the uranium content.

The age A of a sample is given by the relation shown in the equation below, where ρ_s is the number of spontaneous fission tracks per unit area, ρ_i is the number induced per unit area, f is the fraction of the uranium caused to fission by the neutron irradiation, and λ_D and λ_F are the total decay rate and the spontaneous-fission decay rate (decays/atom per unit time) for ^{238}U.

$$\frac{\rho_s}{\rho_i} = \frac{[\exp(\lambda_D A) - 1]\lambda_F}{\lambda_D f}$$

Many minerals and glasses are suitable for fission track dating. To be usable, a sample both must retain tracks over geological time and must contain sufficient uranium so that the track density is high enough to be countable in a reasonable period of time. Thus quartz and olivine are highly track retentive but too low in uranium to be usable, whereas calcite may have enough uranium but appears to be inadequately track retentive. Minerals that have been successfully used for fission track dating include the following: zircon, apatite, allanite, epidote, sphene, hornblende, albite, diopside, enstatite, muscovite, phlogopite, and assorted natural glasses.

Applications. One feature unique to this dating technique is the time span to which it is applicable. It ranges from less than 100 years for certain synthetic, decorative glasses to approximately 4,500,000,000 years, the age of the solar system. This great flexibility in age span allows fission track dating to fill the gap that previously existed between the upper limit of 50,000 years, in which carbon dating is applicable, and the lower limit of roughly 1,000,000 years, below which potassium-argon dating becomes excessively laborious.

A second useful feature is that measurements can sometimes be made on extremely minute specimens, such as chips of meteoritic minerals or fragments of glass from the ocean bottom.

A third useful feature is that each mineral dates the last cooling through the temperature below which tracks are retained permanently. Since this temperature is different for each mineral, it is possible to measure the cooling rate of a rock by dating several minerals—each with a different track-retention temperature. It is similarly possible to detect and date the reheating of an old rock (caused, for example, by a local igneous intrusion or regional metamorphism) by the finding of a low age for the less retentive minerals and a greater age for the more retentive ones. *See* AMINO ACID DATING; EARTH, AGE OF; GEOLOGIC TIME SCALE; MINERAL; ROCK AGE DETERMINATION. R. L. Fleischer; P. B. Price; R. M. Walker

Bibliography. G. Faure, *Principle of Isotope Geology,* 2d ed., 1986; G. A. Wagner and P. Van Den Huate, *Fission Track Dating*, 1992.

FitzGerald-Lorentz contraction

The contraction of a moving body in the direction of its motion. In 1892 G. F. FitzGerald and H. A. Lorentz proposed independently that the failure of the Michelson-Morley experiment to detect an absolute motion of the Earth in space arose from a physical contraction of the interferometer in the direction of the Earth's motion. According to this hypothesis, as formulated more exactly by Albert Einstein in the special theory of relativity, a body in motion with speed v is contracted by a factor $\sqrt{1 - v^2/c^2}$ in the direction of motion, where c is the speed of light. *See* LIGHT; RELATIVITY. E. L. Hill

Flame

An exothermic reaction front or wave in a gaseous medium. Consider a uniform body of gas in which an exothermic chemical reaction (that liberates heat) is initiated by raising the temperature to a sufficiently high level; the reaction is started by a localized release of heat, as by a sufficiently energetic spark, and then spreads from the point of initiation. If the reaction is relatively slow, the whole gas will be involved before the initial region has finished reacting. If, however, the reaction is relatively fast, the reaction zone will develop as a thin front or wave propagating into the unreacted gas, leaving fully reacted gas behind. If the front, in addition, shows luminosity (emission of light), the flame may be considered a classical example. However, perceptible emission of visible radiation is not essential to the definition, and some flames are indeed vanishingly faint under ordinary viewing conditions.

Sufficient reaction rates may also be attained under special conditions (when the gas is very slowly heated inside a closed vessel) without very high temperatures if free radicals are generated in good concentration; this gives so-called cool flames.

Flame-producing reactions. The most common flame-producing reaction is combustion, which is broadly defined as a reaction between fuel and an oxidizer. The oxidizer is typically oxygen (usually in air), but a variety of other substances (for example, bromine with hydrogen) can play the same role in combination with the right fuel. While the overall theoretical reaction in a combustion flame—namely, fuel and oxidizer making fully oxidized products such as carbon dioxide and water vapor—is invariably simple, the actual reaction mechanism is typically very complex, involving many intermediate steps and compounds. Free radicals are generally present and figure prominently in the mechanism. *See* FREE RADICAL.

An overall reaction involving just one reactant is chemical decomposition, for example, ozone decomposing into oxygen. Decomposition flames are usually simpler chemically than combustion flames.

Combustion flames. Combustion flames are broadly divided into premixed flames and diffusion flames. Premixed flames occur when fuel and oxidizer are mixed before they burn. Diffusion flames occur when fuel and oxidizer mix and burn simultaneously. The intermediate case, with partial premixing, has been of relatively low theoretical and practical interest. Flames are further categorized on the basis of shape, time behavior (stationary or moving), flow regime (laminar or turbulent), buoyancy regime (forced convection or natural convection), presence or absence of confinement (as by combustion chamber walls), and flow complications (such as swirling flow and crosswind).

Premixed flames. A flame can be produced in a premixed gas of fuel and oxidizer only if the composition falls between the flammability limits. Once initiated, the flame continually advances into the unburned gas and leaves burned gas behind. The situation is simplest when the unburned gas is still or in laminar motion. Like a sound wave, the flame may be considered to be moving everywhere in the direction normal to its plane. The speed of its advance relative to the unburned gas is the laminar flame speed υ_L (the name and symbol most frequently seen in the older literature are burning velocity and S_u). Away from walls and wherever the flame is thin relative to its curvature, υ_L depends only on the gas properties, and it is this value of υ_L that is usually used. The range of υ_L for near-stoichiometric mixtures of organic fuels with air at 77°F (25°C) and atmospheric pressure is about 0.8–80 in./s (2–200 cm/s), and 16 in./s (40 cm/s) is average. Simplest theory suggests that υ_L is independent of pressure, but the actual behavior is more complex. The effect of increasing the absolute unburned gas temperature T_u is generally to increase υ_L, roughly in proportion to the $^3/_2$ power of T_u.

Figure 1 illustrates some typical features of a premixed laminar flame front. An observer at rest with respect to the flame sees unburned gas (reactants) flowing into the reaction zone from one side and burned gas (products) emerging from the other. The concentrations of fuel and oxidizer fall across the reaction zone, those of intermediate products rise and then fall, and those of final products continually rise. The flame is divided into several zones by the phenomena. The luminous zone is embedded in the reaction zone. The temperature rise accompanying the reaction produces a strong temperature gradient across the flame. Heat is conducted down this gradient, from the departing hot gas into the entering cold gas. The reaction zone is thus preceded by a preheat zone in which the unburned gas is warming but not yet reacting. From this description, it might be assumed that heat conduction into the cold gas, bringing it up to reaction temperature levels, is the controlling mechanism in laminar flame propagation in premixed gases. However, diffusion of free radicals toward the incoming unburned gas also plays a part in kindling reaction and may sometimes, as in cool flames, be dominant.

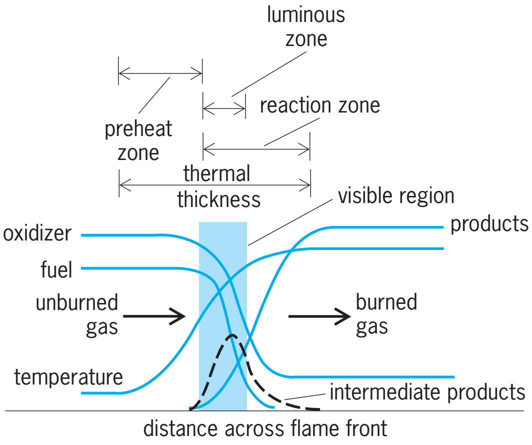


Fig. 1. **Structure of a premixed laminar flame front.**

The thickness of a laminar premixed flame is commonly measured by the thermal thickness—the thickness of the zone of appreciable temperature gradient. In combustion flames of organic fuels with air at atmospheric pressure, cold reactant temperature [77°F (25°C)], and near-stoichiometric composition, the thermal thickness λ is about 0.04–0.2 in. (1–5 mm). It is very nearly inversely proportional to the pressure P, and the product $\upsilon_L \lambda P$ is roughly constant.

The effect of turbulence on flame propagation in premixed gases is complex. A rough approximation is that the flame is locally laminar, but is everywhere more or less strongly convoluted or "wrinkled" by the turbulence. The volume of the reaction zone is thereby increased and, with it, the rate of burning. A common way of characterizing turbulent flame propagation is to define a surface following the average position of the front (smoothing out the wrinkles). The speed of advance of this mean front relative to the mean motion of the unburned gas is called the turbulent flame speed υ_T. The turbulent flame speed is generally greater than the laminar flame speed υ_L, and is a complicated function of υ_L and of the properties of the turbulence in the unburned gas (particularly the spatial scale and the level of velocity fluctuation).

Premixed flames occur in a great variety of situations, both contrived and accidental. In technical applications with continuous combustion, burner flames are most commonly used. The bunsen burner is a familiar example. Typically, premixed gas issues from an orifice or pipe over which the flame hovers in the form of a hollow cone, stabilized at its base on the burner rim. The stabilization mechanism is too complex for simple explanation. The flow of the unburned gas and the propagation of the flame into that gas are exactly balanced, so the flame stands still. When the gas flow is increased, the flame cone lengthens, and when it is decreased, it shortens. When the flow is made too high, stabilization fails and the flame is swept away downstream (blowoff). Stabilization also fails when the flow is made too low, in which case the flame propagates upstream into the burner (flashback). A flame in a stream of premixed gas can also be stabilized in the wake of a nonstreamlined solid body placed in the stream, a so-called bluff-body stabilizer. In applications where flame is initiated in a finite body of gas rather than a continuing stream, as in spark-ignited internal combustion engines, the flame sweeps through the gas from ignition point to boundaries, leaving it more or less completely burned. In engines, high unburned-gas temperatures produced by precompression and high turbulence levels promoted by intake-valve and combustion-chamber design give high effective flame speeds needed for efficient high-speed operation. *See* COMBUSTION CHAMBER; INTERNAL COMBUSTION ENGINE.

Diffusion flames. In technical applications (for example, furnaces, boilers, gas heating devices, and jet and gas-turbine engines), these are typically burner flames designed for continuous combustion, and are much more widely used than premixed flames for such applications. Industrial diffusion flames of any size are also generally turbulent. However, laminar flames do occur in a number of important situations: small gas-jet or vapor-flow flames, as on gas-range burners and candles; the burning of liquid-fuel droplets and metal particles; and the burning of particles or small pieces of solid organic fuels such as wood and coal.

All burning in flames associated with solid and liquid fuels remains in the gas phase. Heat from the flame melts fuels such as candle wax and magnesium metal (at least at the surface) and vaporizes the liquid. The vapor diffuses away and, at the flame front, meets oxygen diffusing toward it from the surrounding gas and burns. Liquid droplets similarly generate vapor, which then burns. In the case of wood and coal, heating at first drives off combustible gases (volatile combustible matter), which burn in a diffusion flame. The subsequent burning of the residual char, however, is exceptional and may involve no flame. The melting and boiling points of carbon are so high that vaporization is normally impossible. Instead, oxygen is usually considered to diffuse to the char surface and there react directly, a mode of reaction called heterogeneous combustion. For some circumstances, though, a mechanism is favored in which carbon dioxide diffuses to the char surface and reacts to produce carbon monoxide. The carbon monoxide then diffuses away, meets incoming oxygen at a flame front, and burns there, forming carbon dioxide, of which half diffuses back to the char to react with it and the other half diffuses away as final product.

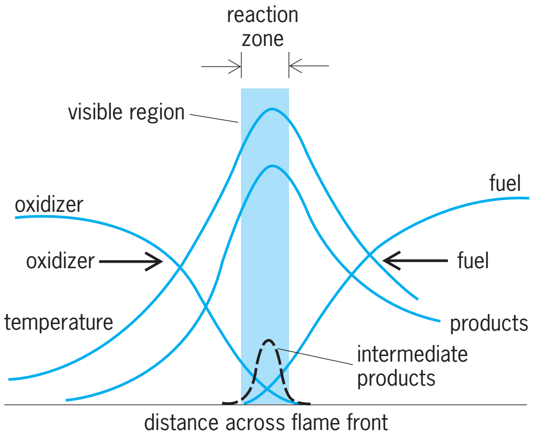


Fig. 2. Structure of a laminar diffusion flame front.

Figure 2 illustrates the structure of a laminar diffusion flame front. Fuel and oxidizer diffuse toward each other and meet and react at the flame front. The reaction zone is the region of overlap between the profiles of fuel and oxidizer, and the temperature peaks in this zone. The reaction zone is usually very thin, which means that the reaction rate of fuel and oxygen is chemical-kinetically very fast, so the fuel and the oxidizer react virtually as fast as they can get together by diffusion. This is why nonpremixed flames are commonly called diffusion flames. Of course, intermediate products of combustion appear in, and tend to be confined to, the reaction zone. Final products are found on both sides of the reaction zone. Any inerts accompanying fuel or oxidizer (such as nitrogen in air) are also found on both sides of the reaction zone. A diffusion flame cannot propagate through the gas, neither into the oxidizer on one side nor into the fuel on the other, but is fixed at the front where fuel and oxidizer meet. The concept of flame speed here is meaningless, and the situation is altogether different from a premixed flame.

The effect of turbulence bears a vague similarity to that in the premixed flame in that here, too, turbulence produces a distorted and greatly convoluted reaction zone. However, turbulence affects the meeting of fuel and oxidizer quite profoundly. A crude way of explaining this is to say that turbulent diffusion is added to, and usually completely overshadows, molecular diffusion. *See* CHEMICAL DYNAMICS; EXPLOSIVE; OXYGEN.

Henry A. Becker

Bibliography. N. A. Chigier (ed.), *Progress in Energy and Combustion Science*, vol. 10, 1986; Combustion Institute, *Combustion Flame*, vol. 55, 1984; *Combustion Science and Technology* (journal), vol. 35, 1984; F. M. Fristrom, *Flame Structure*

and Processes, 1994; I. Glassman (ed.), *Combustion*, 3d ed., 1996; K. Kuo, *Fundamentals of Combustion*.

Flameproofing

The process of treating materials so that they will not support combustion. Although cellulosic materials such as paper, fiberboard, textiles, and wood products cannot be treated so that they will not be destroyed by long exposure to fire, they can be treated to retard the spreading of fire and to be self-extinguishing after the igniting condition has been removed.

Treatments. Numerous methods have been proposed for flameproofing cellulosic products. One of the simplest and most commonly used for paper and wood products is impregnation with various soluble salts, such as ammonium sulfate, ammonium phosphate, ammonium sulfamate, borax, and boric acid. Special formulations are often used to minimize the effects of these treatments on the color, softness, strength, permanence, or other qualities of the paper.

For some applications, these treatments are not suitable because the salts remain soluble and leach out easily on exposure to water. A limited degree of resistance to leaching can be achieved by the addition of latex, lacquers, or waterproofing agents. In some cases the flameproofing agent can be given some resistance to leaching by causing it to react with the cellulose fiber (for example, urea and ammonium phosphate).

Leach-resistant flameproofing may also be obtained by incorporating insoluble retardants in the paper during manufacture, by application of insoluble materials in the form of emulsions, dispersions, or organic solutions, or by precipitation on, or reaction with, the fibers in multiple-bath treatments. The materials involved are of the same general types as those used for flameproofing textiles and include metallic oxides or sulfides and halogenated organic compounds. Because of the higher material cost and more elaborate equipment required, such treatments are used on paper only when unusual conditions justify their cost.

Action of retardants. Various theories have been proposed to explain flameproofing action. In general, these are similar to those considered in the flameproofing of textiles. Two types of combustion may be involved. In one, the volatile decomposition products burn with a flame; in the other, the solid material undergoes flameless combustion, or afterglow. In general, the alkaline types of retardants are effective in preventing afterflame, and the acid types control afterglow. A few, such as ammonium phosphates and halogenated products, reduce both flame and glow.

The theory for prevention of afterflame having the most support is a chemical one, which holds that effective flame retardants direct the decomposition of the cellulose when heated to minimize the formation of volatile flammable products and increase the amount of water and solid char formed. The prevention of afterglow is usually attributed to a modification of the flameless combustion to make it less exothermic and therefore incapable of maintaining itself, for example, by formation of carbon monoxide instead of carbon dioxide. Other theories include the formation of a coating or froth which excludes oxygen and smothers the combustion, the formation of nonflammable gases which dilute the flammable products and exclude oxygen, and the thermal theories which hold that heat is dissipated by endothermic reactions or by conduction. It is probably that these actions do contribute in some cases.

The susceptibility of paper or fiberboard to combustion may also be modified by the introduction of noncombustible but otherwise inert materials, such as mineral fillers, asbestos, and glass or ceramic fibers, into the paper or board during manufacture. The high proportions required result in products with properties determined largely by noncombustible components rather than by cellulosic fibers.

The behavior of paper, fiberboard, or wood products exposed to fire depends also on structural factors. The density of the material and the ratio of mass to surface exposed affect the ease of ignition and the rate of spreading or penetration by the fire. Fire-resistant coatings minimize the spread of flame. Heat-reflective coatings or insulating coatings prevent or retard the material beneath them from reaching the ignition temperature. Intumescent coatings are formulated to form a thick insulating foam when exposed to fire.

A wide variation exists in methods for evaluating the efficiency of a flameproofing treatment or of suitability for a given use. Paper is usually tested by igniting with a flame of specified character for a predetermined period of time and then noting duration of afterflame and of afterglow and the amount of char length or char area. Structural materials are tested by other methods which involve features of construction in addition to those of composition. Methods have been standardized by the Technical Association of the Pulp and Paper Industry (TAPPI), American Society for Testing and Materials (ASTM), and various governmental agencies.

For a discussion of textile flameproofing *see* COMBUSTION; FIRE EXTINGUISHER; FLAME; TEXTILE CHEMISTRY.

T. A. Howells

Bibliography. Fire Retardant Chemicals Association, *Flame Retardants 101: Basic Dynamics*, 1998; A. R. Horrocks and D. Price, *Fire Retardant Materials*, 2001; National Research Council, *Toxicological Risks of Selected Flame-Retardant Chemicals*, 2000.

Flash welding

A form of resistance welding that is used for mass production. The welding circuit consists of a low-voltage, high-current energy source (usually a

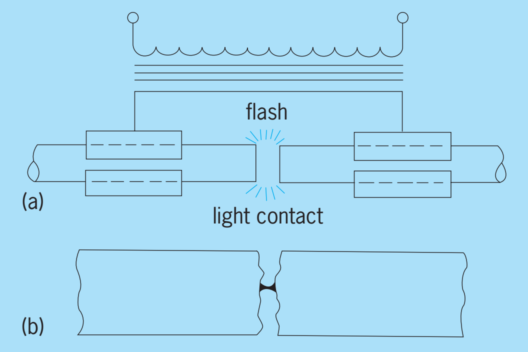


Fig. 1. Elements of flash-welding operation. (*a*) Two pieces of horizontally clamped metal are brought together to make a light contact. At that time the welding circuit is closed to the transformer. (*b*) Current starts flowing when small areas of the workpiece make initial contact. (*After A. L. Phillips, ed., Welding Handbook, 4th ed, American Welding Society, 1957*)

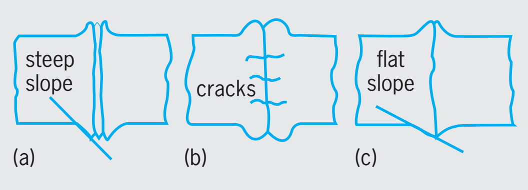


Fig. 2. Flash-weld quality indicated on the surface. (*a*) Satisfactory heat and upset. (*b*) Cracks caused by insufficient heat. (*c*) Insufficient heat or upset force, or both. (*After A. L. Phillips, ed., Welding Handbook, 4th ed., American Welding Society, 1957*)

welding transformer) and two clamping electrodes, one stationary and one movable.

The two pieces of metal to be welded are clamped tightly in the electrodes, and one is moved toward the other until they meet, making light contact (**Fig. 1***a*). Energizing the transformer causes a high-density current to flow through small areas that are in contact with each other (Fig. 1*b*). Flashing starts, and the movable workpiece must accelerate at the proper rate to maintain an increasing flashing action. After a proper heat gradient has been established on the two edges to be welded, an upset force is suddenly applied to complete the weld. This upset force extrudes slag, oxides, molten metal, and some of the softer plastic metal, making a weld in the colder zone of the heated metal (**Fig. 2**). Flash and upset material are frequently removed, although when the design permits, a portion of the upset material may be retained as a reinforcement. *See* ARC WELDING; RESISTANCE WELDING.

Eugene J. Limpel

Bibliography. R. Carr and R. O'Con, *Welding Practices and Procedures*, 1983; H. B. Cary, *Modern Welding Technology*, 4th ed., 1997; J. W. Giachino and W. R. Weeks, *Welding Skills and Practices*, 1985.

Flat-panel display device

An electronic display in which a large orthogonal array of display elements, such as liquid-crystal or electroluminescent elements, form a flat screen.

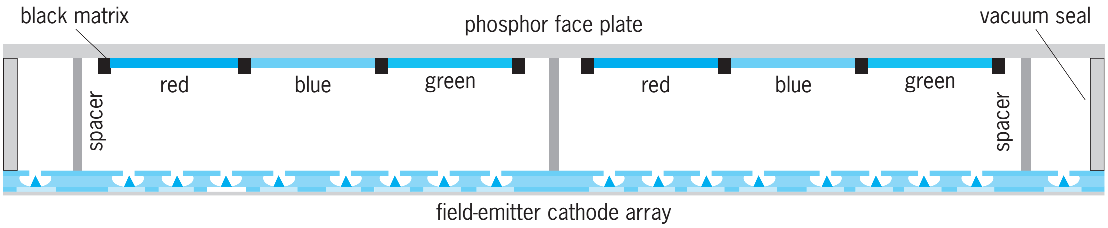
Since the days of the first television tube, display engineers have dreamed of producing flat-panel displays (FPDs) with the visual attributes of a cathode-ray tube (CRT). The term "flat-panel display" is actually a misnomer, since thinness is the distinguishing characteristic. Most television sets and computer monitors currently employ cathode-ray tubes. Cathode-ray tubes cannot be thin because the light is generated by the process of cathodoluminescence whereby a high-energy electron beam is scanned across a screen covered with an inorganic phosphor, one point, or pixel, at a time, activating light-emitting centers in the material. The cathode-ray tube must have moderate depth to allow the electron beam to be magnetically or electrostatically scanned across the entire screen. *See* CATHODE-RAY TUBE; CATHODOLUMINESCENCE; ELECTRONIC DISPLAY.

For a flat-panel display technology to be successful, it must at least match the basic performance of a cathode-ray tube by having (1) full color, (2) full gray scale, (3) high efficiency and brightness, (4) the ability to display full-motion video, (5) wide viewing angle, and (6) wide range of operating conditions. Flat-panel displays should also provide the following benefits: (1) thinness and light weight, (2) good linearity, (3) insensitivity to magnetic fields, and (4) no x-ray generation. These four attributes are not possible in a cathode-ray tube.

Flat-panel displays can be divided into three types: transmissive, emissive, and reflective. A transmissive display has a backlight, with the image being formed by a spatial light modulator. A transmissive display is typically low in power efficiency; the user sees only a small fraction of the light from the backlight. An emissive display generates light only at pixels that are turned on. Emissive displays should be more efficient than transmissive displays, but due to low efficiency in the light generation process most emissive and transmissive flat panel displays have comparable efficiency. Reflective displays, which reflect ambient light, are most efficient. They are particularly good where ambient light is very bright, such as direct sunlight. They do not work well in low-light environments.

Most flat-panel displays are addressed as an X-Y matrix, the intersection of the row and column defining an individual pixel (see **illus.**). Matrix addressing provides the potential for an all-digital display. Currently available flat-panel display devices range from 1.25-cm (0.5-in.) diagonal displays used in head-mounted systems to 125-cm (50-in.) diagonal plasma displays.

Currently, most commercially manufactured flat-panel display devices are liquid-crystal displays (LCDs). The benchmark for flat-panel display performance is the active matrix liquid-crystal display (AMLCD). Most portable computers use AMLCDs. Competing flat-panel display technologies include electroluminescent displays, plasma display panels, vacuum fluorescent displays, and field-emission displays. Electroluminescent displays are often used in industrial and medical applications because of their ruggedness and wide range of operating



Field-emission display, an example of a matrix-addressed display. Each pixel is addressed in an X-Y matrix. In the case of a color display, each pixel is subdivided into a red, a blue, and a green subpixel.

temperatures. Plasma display panels are most often seen as large flat televisions, while vacuum fluorescent displays are used in applications where the information content is fairly low, such as the displays on appliances or in automobiles. Field-emission displays are the most recent of these flat-panel technologies. There are also emerging flat-panel display technologies based on laser-pumped-polymer–switched matrices and electrophoretically switched reflective displays. They offer the potential to provide low-cost, very large displays. Another area of active research in flat-panel displays is the development of flexible displays produced on plastic substrates, which have the potential for a tremendous number of applications.

Liquid-crystal displays. The liquid-crystal display is a spatially addressed light valve that transmits or reflects light of a given polarization, depending on the orientation of the liquid crystal molecule. The orientation of the molecule is determined by how the material is affixed to the substrate (the pre-tilt angle) and the applied electric field placed across the molecule. There are several variations of the liquid-crystal display, including the twisted nematic, the ferroelectric, the polymer dispersed, and the cholesteric liquid-crystal display. Each of these technologies has specific strengths and weaknesses, ranging from high-speed switching for ferroelectric liquid crystals to bistable operation for cholesteric liquid crystals. Advanced liquid-crystal displays sometimes include active matrix addressing. In an active matrix addressing scheme, one or more thin-film transistors are placed at each pixel to more accurately control the applied voltage on the pixel, thus increasing both the refresh rate and the dynamic range of gray scale. Amorphous silicon is used in the manufacture of thin-film transistors; for high-performance applications, polycrystalline silicon is used.

Liquid-crystal displays can operate in either a transmissive or a reflective mode. In the transmissive mode, a backlight and a polarizer are required. To produce a color display, each pixel is usually divided into at least three subpixels, with a red, blue, or green filter placed above each subpixel. In the reflective mode, ambient light is reflected from the liquid crystal by using a reflective material on the backside of the liquid-crystal display, or the liquid-crystal material itself can be reflective. The deficiencies of current liquid-crystal displays include (1) limited viewing angle for both intensity and color purity, (2) limited operating temperature range, (3) low efficiency for light utilization (50% polarization loss, 66% loss for color filters), and (4) slow refresh rates for reflective displays. These deficiencies are being reduced with continuing research. Advanced AMLCDs exhibit resolutions greater than 140 lines per inch (color; 55 lines per centimeter) and diagonal dimensions as large as 30 in. (76 cm). *See* LIQUID CRYSTALS.

Electroluminescent displays. Electroluminescent (EL) displays are emissive displays in which light is generated by the passage of current through an electroluminescent layer. Traditionally they have been based on inorganic materials, but recently there has been significant interest in organic-based light-emitting materials. In inorganic electroluminescent displays the light is generated through impact ionization of the active sites in the phosphor layer by energetic electrons. In organic-based light-emitting diodes, light is generated through electron-hole pair recombination. Inorganic displays are high-voltage devices (more than 100 V), while organic displays are generally low voltage (less than 10 V). The electroluminescent process is typically not as efficient in generating light as the cathodoluminescent process used in cathode-ray tubes. A main advantage of the electroluminescent display is that it is completely solid-state. Large-area [greater than 3-in. (7.6-cm) diagonal] electroluminescent displays are normally matrix-addressed devices, similar to a passive matrix liquid-crystal display. Recently, active matrix addressing has been introduced to small diagonal [less than 5 cm (2 in.)] electroluminescent displays, where the display is produced directly on a silicon substrate. Active matrix addressing provides faster displays with larger dynamic range, as well as higher brightness due to higher-frequency operation. Also, active matrix addressing reduces the interconnect density for high-resolution displays. Active matrix electroluminescent displays with 12-micrometer-size pixels [2000 lines per inch (800 lines per centimeter)] are currently under development for small-format displays. *See* ELECTROLUMINESCENCE; LIGHT-EMITTING DIODE.

Plasma display panels. The plasma display panel (PDP) is an emissive display in which light is generated via an excited plasma in a low-pressure gas. In a color plasma display panel, the plasma generates deep ultraviolet photons, which in turn excite photoluminescent phosphors to produce the colors for the display. Resolution is limited by the physical barriers required to separate the plasma in the individual subpixels. Plasma display panels are typically low- to moderate-resolution devices that have relatively poor power efficiency. They are best suited for large, low- or moderate-resolution display applications such as wall television. Plasma display panels generally are not suitable for portable applications. There have been significant improvements in the power efficiency and the lifetime of color plasma panels. Plasma panels as large as 125-cm (50-in.) diagonal are in production, and prototype panels with a resolution of 80 lines per inch (32 lines per centimeter) are being produced.

Vacuum fluorescent displays. The vacuum fluorescent display is based on the thermionic emission of electrons from a small wire, and the cathodoluminescent production of light. Most vacuum fluorescent displays operate at low anode voltage, typically less than 50 V, and use low-voltage phosphors such as zinc oxide. The display is formed by segmenting the phosphor and applying the anode bias to the segments to be illuminated. The resolution is moderate to low, such as for small alphanumeric displays. The main advantages of the vacuum fluorescent display are high brightness, low operating voltage, and a wide operating temperature range, making it attractive for automotive applications.

Field-emission displays. A field-emission display (FED) is similar to a cathode-ray tube in many ways. The major difference is the type and the number of electron emitters. The electron guns in a cathode-ray tube produce electrons by thermionic emission from a cathode. CRTs typically have one or three electron guns, depending on the configuration of the electron scanning system. The extracted electrons are focused by the electron gun, and while the electrons are accelerated toward the anode plate, electromagnetic deflection coils are used to scan the electron beam across a phosphor-coated faceplate. This requires a large distance between the deflection coils and faceplate. The larger the cathode-ray tube viewing area, the greater the depth (anode to cathode distance) required to scan the beam.

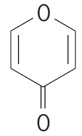
By contrast, in a field-emission display, electrons are produced by cold cathode emission. Several electron emitters or cathodes are associated with each pixel (see illus.). This eliminates the need for beam steering. Field-emission of electrons is achieved by applying a bias of less than 100 V between the cathode and a control electrode. These extracted electrons are then accelerated without deflection toward a phosphor-coated flat faceplate, biased between 300 and 10,000 V. Unlike the cathode-ray tube, electrons in a field-emission triode (cathode-gate-anode) device receive little focus. Thus, in a field-emission display, in order to achieve a small spot size, a small gap (less than 1 mm) is maintained between the cathode array and the anode faceplate. As a result, a field-emission display is much thinner than a cathode-ray tube. Also, since the emitters are placed in an addressable X-Y matrix, a field-emission display does not suffer from the traditional nonlinearity and pin cushion effects associated with a cathode-ray tube. A field-emission display, therefore, combines the best features of a cathode-ray tube with the thinness and linearity of a flat-panel display. *See* FIELD EMISSION.

Bruce E. Gnade

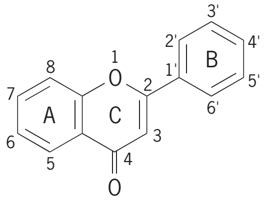
Bibliography. B. R. Chalamala, Y. Wei, and B. Gnade, FED up with flat tubes, *IEEE Spectrum*, 35:42, 1998; G. Gu and S. Forrest, Design of flat-panel displays based on organic light emitting devices, *IEEE J. Selected Top. Quant. Electr.*, 4:83, 1998; Liquid crystal devices display their empire, *Electr. Des.*, 45:40, 1997; L. W. MacDonald and A. C. Lowe (eds.), *Display Systems: Design and Applications*, John Wiley, New York, 1997; L. E. Tannas, Jr. (ed.), *Flat-Panel Displays and CRTs*, Van Nostrand Reinhold, New York, 1985; K. Werner, The flowering of flat displays, *IEEE Spectrum*, 34(5):40–49, May 1997.

Flavonoid

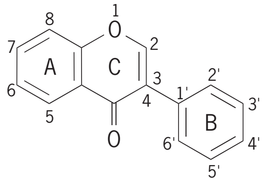
A large category of natural plant products that derive from γ-pyrone (**1**). All flavonoid compounds, which are derived from either 2-phenylbenzopyrone (**2**) or 3-phenylbenzopyrone (**3**), can be classified into 10 groups (**Fig. 1**): chalcones, flavanones, flavones, flavonols, anthocyanidins (flavylium cations), flavan 3-ols (catechins), flavan 3,4-diols (proanthocyanidins), biflavonoids and oligomeric flavonoids, isoflavonoids, and the aurones. They differ in the oxidation level or substitution pattern of their heterocyclic ring (ring C).



(**1**)



(**2**)



(**3**)

Physical and chemical properties. Most flavonoids show characteristic absorbance in the visible and ultraviolet region of the spectrum. Depending on the basic flavonoid skeleton, the visible absorption bands may be located between 330 and 370 nanometers. The absorption band in the ultraviolet region is less affected by structural variations and is usually between 260 and 280 nm. Because of the close

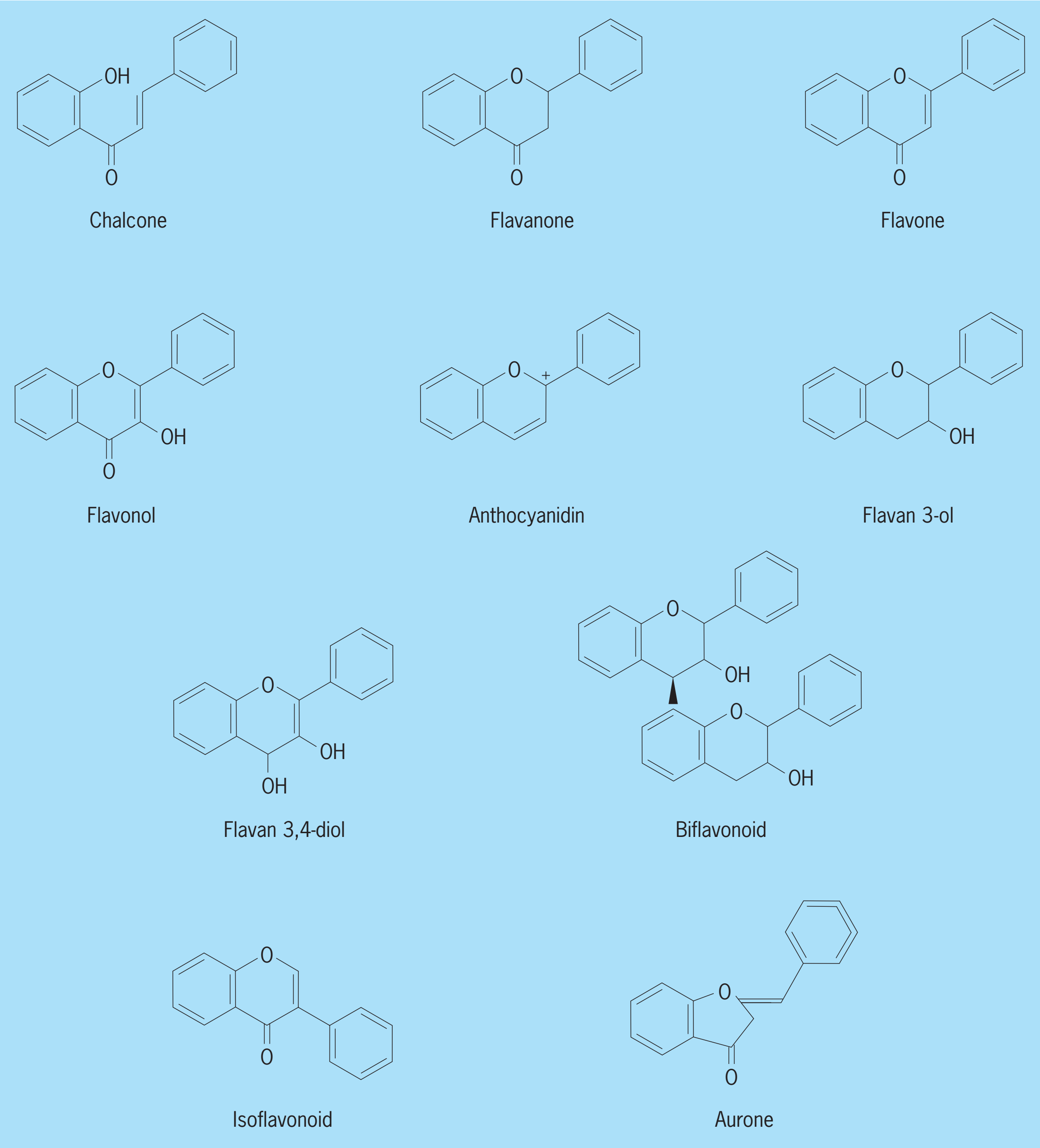


Fig. 1. The 10 classes of flavonoid compounds and their structural relationship.

relationship between the positions of these absorption bands and the basic structure of the flavonoid skeleton, they have proven essential in structural determinations.

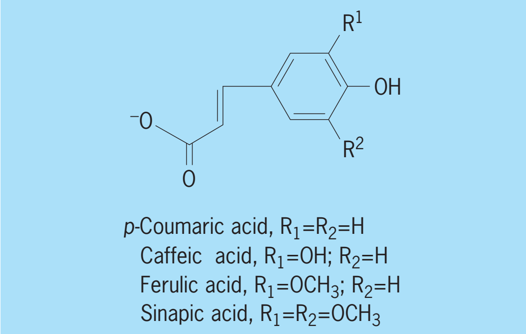


Fig. 2. Aromatic acids that may be attached to the sugar moiety of the flavonoid compounds.

More than 1300 different flavonoid compounds have been isolated from plants. Whereas the basic differences between the various classes of flavonoids lie with the oxidation stage of the heterocyclic ring and the site of attachment of the phenyl ring (ring B) to the heterocycle, individual flavonoids in a group differ from each other by the number and position of the hydroxy, methoxy, and sugar substituents. Thus positions 3′, 4′, 5, and 7 (structures 2 and 3) are the most commonly substituted among the flavonones and flavones, and 3, 3′, 4′, 5, and 7 among the flavonols, anthocyanidins, and flavan 3-ols. The number of the substituents may vary from zero to eight, but a combination of four to six hydroxyl and methoxyl groups is found most frequently. As a rule, flavonoid compounds occur in plants as glycosides, with hexoses such as glucose, galactose, and rhamnose, and pentoses such as arabinose and xylose as the most commonly found sugars. The sugars can be

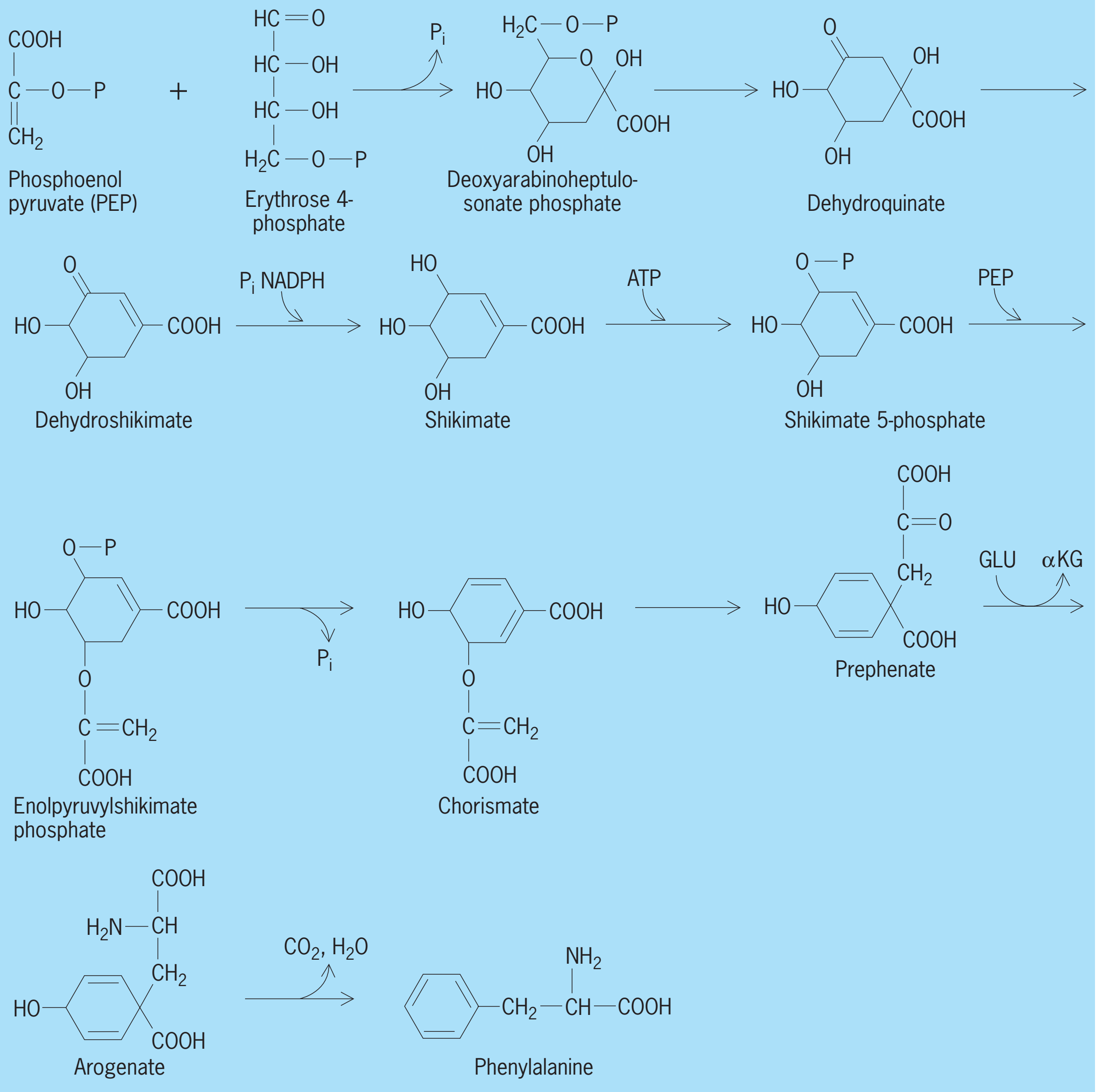


Fig. 3. Shikimate section of the plant aromatic pathway that produces phenylalanine (and the other aromatic amino acids tyrosine and tryptophan). P_i = inorganic phosphate; GLU = glutamate; αKG = α-ketoglutarate.

attached singly or in combination with each other. The most common position of glycosylation of the flavanones and flavones is the 7-hydroxyl, in flavonols the 3- and the 7-hydroxyls, and in anthocyanins the 3- and 5-hydroxyls. Glycosylation renders these compounds water-soluble and permits their accumulation in the vacuoles of cells. *See* GLYCOSIDE.

The sugars in the molecules often have aliphatic (acetic, malonic) or aromatic (*p*-coumaric, caffeic, ferulic, and sinapic) acids attached (**Fig. 2**). The aliphatic acids are difficult to detect by spectral investigation, but the aromatic acids are easily found by the appearance of a new band between 307 and 340 nm in the spectra. Sugars may also be attached by carbon-carbon (C-C) bonds to flavonoids (C-glycosides); the most commonly found attachment sites are positions 6 and 8. Flavonoids that carry terpenoid or alkaloid side chains are known to exist, but their occurrence is rare. Chalcones and aurones are intensely yellow and are the cause of

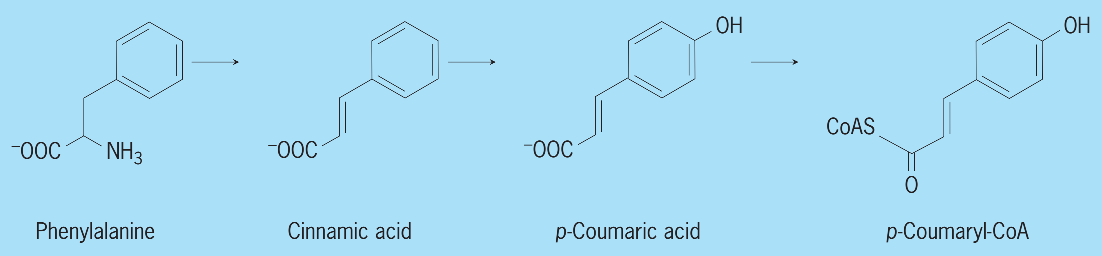


Fig. 4. Phenylpropanoid section of the plant aromatic pathway.

flower color in some plants (chalcones in carnations and Dahlia species and aurones in the daisy family).

Biosynthesis. As complicated as the structures of the diverse flavonoids may appear, their biosynthesis starts from simple components. The pathway that leads to their formation, referred to as plant aromatic metabolism, is one of the best understood in plants. It comprises the shikimate pathway, the phenylproponoid segment, and the flavonoid segment. The shikimate pathway produces the aromatic amino acids phenylalanine, tyrosine, and tryptophan; the phenylpropanoid segment transforms phenylalanine to cinnamic acid derivatives; and the flavonoid segment condenses the activated cinnamic acid derivative, *p*-coumaryl-CoA, with three molecules of malonyl-CoA to form the first flavonoid compound, naringenin chalcone, from which all flavonoids derive. *See* AMINO ACIDS.

The shikimate section of the pathway starts with a simple sugar derivative, erythrose 4-phosphate (**Fig. 3**); its condensation with another simple compound, phosphoenol pyruvate, yields a seven-membered sugar, deoxyarabinoheptulosonate phosphate. That sugar cyclizes and undergoes a series of transformation reactions to form shikimate phosphate, for which the pathway segment is named. Shikimate phosphate then undergoes a further condensation reaction with phosphoenol pyruvate, rearrangement of this newly added side chain, amination, and in the case of phenylalanine, a loss of water.

One part of the phenylalanine so formed is deaminated by the first enzyme of the phenylpropanoid pathway segment, phenylalanine ammonia lyase, to form cinnamic acid (**Fig. 4**). Introduction of a hydroxyl group in the para position and condensation with coenzyme A produces *p*-coumaryl-CoA, the activated form of the phenylpropanoids. That may be further hydroxylated and methoxylated to the aromatic acid derivatives shown in Fig. 2 and may participate in the production of the plant structural component lignin. *See* LIGNIN.

Condensation of the activated *p*-coumaric acid with three similarly activated malonic acids by the enzyme chalcone synthase produces the first flavonoid compound, naringenin chalcone, from which all flavonoids derive. The flavonoid segment of the plant aromatic metabolism is shown in **Fig. 5**, using the example of kaempferol 3-glucoside, the simplest flavonol derivative in plants. A series of modifications of naringenin chalcone leads to the formation of all flavonoid classes and their glycosylated forms that occur naturally in plants. From an evolutionary point of view, the first enzymatic step of the flavonoid pathway section shows the condensation of three malonyl CoA's with the starter compound *p*-coumaryl-CoA to resemble in its mechanism that of the fatty acid synthesis that uses acetyl-CoA as a starter. This similarity with the otherwise different fatty acid metabolism suggests that flavonoid and fatty acid metabolism may have had common ancestral origins.

The biosynthetic mechanism of the isoflavonoids has two unique reactions not found in the pathway segments of other flavonoids. One is the formation of 5-deoxy isoflavonoids in legumes, wherein a specific reductase coacts with chalcone synthase to remove the hydroxyl prior to cyclization of the A ring. The other reaction is the migration of the aromatic B ring from the 2 to the 3 position. This reaction is catalyzed by the enzyme isoflavone synthase, a monooxygenase that requires reduced nicotinamide adenine dinucleotide phosphate, and is believed to take place via an enol-epoxide intermediate.

Controls over biosynthesis. In order to maintain a cell's metabolism and prevent overproduction of one or the other metabolite, biochemical events have to be strictly regulated. This is generally accomplished in cells and organisms by a mechanism called feedback control. In this mechanism, overproduction of

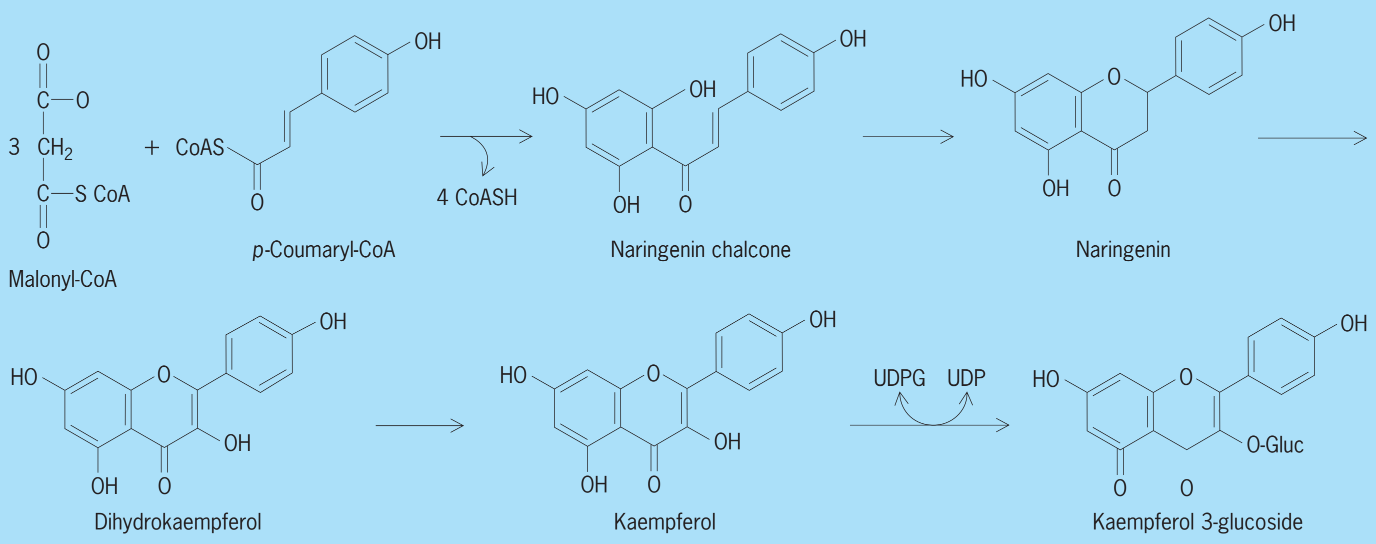


Fig. 5. Flavonoid section of the plant aromatic pathway. UDPG = uridine-5′-diphosphoglucose; UDP = uridine-5′-diphosphate.

one or more compounds has an inhibitory effect on enzymes located further upstream in the pathway, shutting them down until the concentration of the inhibitory product diminishes. Such a feedback control has been shown to exist in the shikimate segment of the flavonoid pathway, but no such control has been found with the other two segments. They are controlled at the genetic level by de novo synthesis of the enzymes from their corresponding messenger ribonucleic acids (**Fig. 6**). The pathway segments are thought to function as enzyme complexes in which the product of the one enzyme is used as substrate by the next until the final product is formed.

Sites of accumulation and synthesis. There are few data on cell or tissue-specific accumulation of flavonoids, primarily because of the difficulties encountered in isolating individual tissues that are free of cross contamination. The few reports available indicate that flavonoids accumulate in epidermal tissues, with approximately 70% in the upper and 30% in the lower epidermis. The development of methods to isolate intact, mature plant vacuoles on a large scale led to investigations that showed the subcellular accumulation site of flavonoids to be vacuolar.

Although the vacuoles are probably the only site of flavonoid accumulation in the cells, they are clearly not the site of synthesis. Instead, synthesis of flavonoids takes place in the cytoplasm, and evidence indicates that the different pathway segments may function as membrane-associated enzyme complexes.

Physiological role in plants. Flavonoid compounds can no longer be regarded as stray end products of metabolism; some are now known to be physiologically active. The symbiotic interaction of *Rhizobium meliloti* and some members of the family Leguminosae results in the formation of nitrogen-fixing root nodules. Specific genes are required for the early host responses of cortical cell division and root hair curling. Induction of this gene expression was found to be caused by host-symbiont signaling, and a number of flavonoid compounds were discovered to be the host-specific signal molecules. In alfalfa, the specific host signal is the 3′,4′,5,7-tetrahydroxyflavone, luteolin (**4**). *Trifolium* species produces 4′,7-dihydroxyflavone (**5**), 4′,7-dihydroxy-3′-methoxyflavone, and 4′-hydroxy-7-methoxyflavone as the specific host signals for nodulation. *Vicia sativa* synthesizes a number of nodulation-inducing compounds; at least one was reported to be a flavonoid. *See* NITROGEN FIXATION.

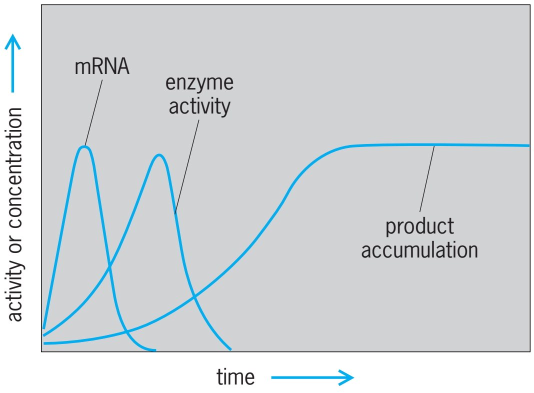
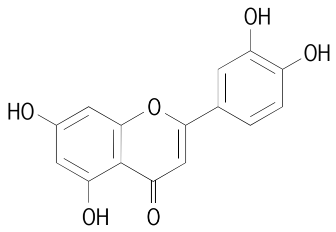
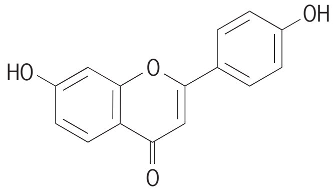


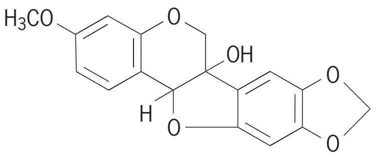
Fig. 6. Control of flavonoid synthesis in plants. A signal (either light or stress) induces messenger ribonucleic acid (mRNA) production, which synthesizes the corresponding enzyme. The enzyme in turn synthesizes the product, which accumulates when the pathway is blocked, or acts as the substrate for the next enzyme, which is concomitantly induced.



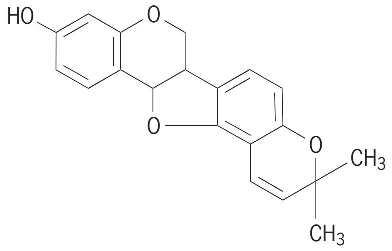
(4)



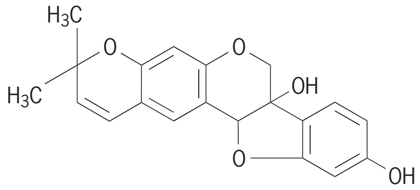
(5)



(6)



(7)



(8)

Light is an essential energy source for plants, and it provides external signals that regulate the developmental processes and adjustments to changes in the environment. In tissues exposed to potentially

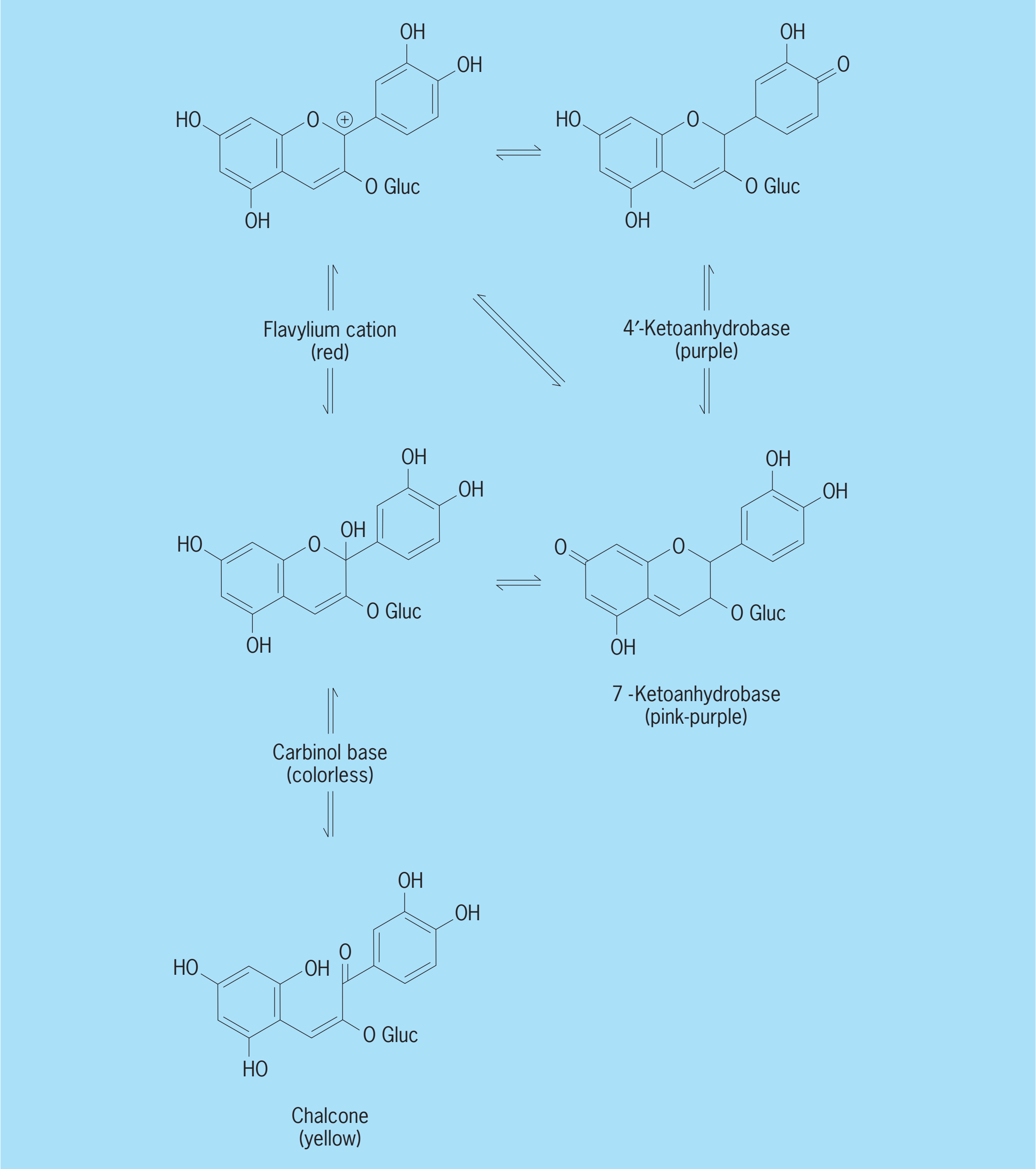


Fig. 7. Structural transformation reactions of anthocyanins as shown by the example of cyanidin 3-glucoside, in the pH range 1–7, which contributes to color variation in plants. O Gluc = O glucose.

damaging ultraviolet radiation, a major form of adjustment is the rapid accumulation of ultraviolet-blocking compounds. The enzymatic machinery for flavonoid production is induced by ultraviolet irradiation. Flavonoids accumulate in the vacuoles of epidermal cells and absorb strongly in the critical range of 280–380 nm, where damage caused by ultraviolet radiation occurs. Flavonoids have therefore been linked to the mechanism of plants for protection from ultraviolet radiation.

Many plant species synthesize phytoalexins upon invasion by microorganisms. The production of these antimicrobial compounds represents an active means of defense against potential pathogens. The majority of phytoalexins produced by legumes are isoflavonoids, and each plant species seems to produce a specific compound. Thus peas (*Pisum sativum*) produce pisatin (**6**), beans (*Phaseolus vulgaris*) produce phaseollin (**7**), and soybeans (*Glycine max*) produce glyceollins (**8**) as part of their defense mechanism.

Anthocyanins. Because of their strikingly vivid color, ranging from deep red through purple to deep blue, anthocyanins represent the most visible class of flavonoid compounds. Cyanidin is probably the most widely distributed anthocyanin. Other common anthocyanidins are pelargonidin, peonidin, delphinidin, petunidin, and malvidin, all of which occur in plants as glycosides. Anthocyanins, the glycosides of anthocyanidins, are most obvious in flowers and fruits, but they are also present in roots, stems, leaves, seeds, and other parts of the plant. In autumn, when the cool nights slow down respiration, photosynthate accumulates, chlorophyll breaks down, and anthocyanins accumulate in the leaves. The accumulated anthocyanins, together with carotenes, provide the varied colors characteristic of that season. Anthocyanins are also produced when plants are subjected to other stress, such as ultraviolet radiation, injury by insects, malnutrition, or unusual concentrations of metal ions.

Color variation in plants results from many factors. Variation in color intensity may be caused by differences in concentration of the same pigment or by variations in the acidity of the cell sap. Anthocyanins undergo dramatic color changes when the acidity of their solution is varied, a response caused by reversible structural transformation reactions (**Fig. 7**).

In the pH range of 1 and below, all anthocyanin molecules of a solution are present in the flavylium cation form, which is red. The color of the flavylium cation is caused by a positive charge in the 2 position of the molecule and the double bond between carbons 3 and 4, which permits electron flow between the two aromatic rings A and B. As the pH of an anthocyanin solution is raised to between 4 and 4.5, the flavylium cation may undergo two different reactions: the nucleophilic attack of water in the 2 position, which yields the colorless carbinol base; and the relocation of electrons from ring B or A to produce the corresponding anhydro bases with the 4′-keto or 7-keto structures. Both have a purplish color that increases in stability as the pH rises toward neutrality. The anhydrobases also undergo hydration, which results in the formation of the colorless carbinol base. Around neutrality, the carbinol base undergoes ring fission to produce the yellow chalcone. All such structural transformation reactions are freely reversible, and the red color of an anthocyanin solution can be reestablished fully upon readjusting the pH to 1 or below.

Flower petals exhibit great color variation that ranges from pink to blue. Whereas in old, senescing petals the blue color may be due to a rise of pH in the vacuole, the color of most flowers is caused by complex formation of the anthocyanins with other phenolic, or flavonoid, compounds.

The varied colors of fruits, on the other hand, are not caused by complexes with other compounds but by variation of the pigment concentration in the tissues and cells, and the pH of the cell sap. Thus the blue color of plums or some grapes is not caused by a blue pigment in the cells but by a combination of a very high concentration of anthocyanins that reflects light and a waxy layer that helps to diffuse it. *See* PHYTOCHROME; PLANT METABOLISM.

Geza Hrazdina

Bibliography. M. Dean, *Naturally Occurring Oxygen Ring Compounds,* 1963; J. B. Harborne, *The Flavonoids: Advances in Research since 1980*, 1988; J. B. Harborne and T. J. Mabry (eds.), *The Flavonoids*: *Recent Advances in Research*, 1982; C. F. Van Sumere and P. J. Lea (eds.), *The Biochemistry of Plant Phenolics*, 1985.

Flavor

Any of the six different varieties of quarks. All hadronic matter is composed of quarks, the most elementary constituents of matter. The six different flavors are labeled u, d, s, c, b, and t, corresponding to up, down, strange, charmed, bottom, and top. Quarks are all spin-$^1/_2$ fermions. The u, c, and t flavors carry a positive electric charge equal in magnitude to two-thirds that of the electron; the d, s, and b flavors have a negative charge one-third that of the electron. Different flavored quarks have vastly different masses ranging from the lightest, the u quark, with a mass around 5 MeV/c^2 (where c is the speed of light), equal to the mass of about 10 electrons, to the top quark, with a mass 35,000 times greater, or 175 GeV/c^2, about the mass of a gold atom. Because of its large mass, only in 1994 was evidence presented for existence of the top quark. Quarks of any flavor are further characterized by three additional quantum numbers called color: red, green, and blue. Each quark has an antiquark counterpart carrying the corresponding anticolor. All of these properties have been well established by experiment. *See* COLOR (QUANTUM MECHANICS).

Physically realizable particles are colorless. For example, the proton, composed of two u- and one d-flavored quark, has one red, one green, and one blue quark to make the ensemble colorless. The fact that quarks carry color and physical entities must be colorless accounts for their never appearing as isolated objects. Mesons are composed of quark-antiquark $(q,\mathbf{q}^{-})$ combinations. A positively charged pion consists of a u and anti-d quark; the anticolor of the anti-d quark cancels the color of the u quark. *See* BARYON; MESON; PROTON.

The strong interquark forces are produced by the exchange of massless, spin-1 bosons called gluons. All evidence suggests that these forces are flavor independent despite the vastly different masses of quarks of different flavors. Furthermore, electromagnetic forces are also flavor independent. In contrast, the weak interactions can induce a change in flavor. For example, at the quark level the beta decay of the neutron is viewed as the decay of the d quark to a u quark plus an electron and a neutrino. *See* GLUONS.

The six different quark flavors are frequently grouped with the leptons into three so-called generations, as shown below, where their electrical

$$\begin{bmatrix} u \\ d \\ e \\ \nu_e \end{bmatrix} \begin{bmatrix} c \\ s \\ \mu \\ \nu_\mu \end{bmatrix} \begin{bmatrix} t \\ b \\ \tau \\ \nu_\tau \end{bmatrix} \quad \begin{matrix} {}^2/_3 \\ -{}^1/_3 \\ -1 \\ 0 \end{matrix}$$

charges, row by row, are given on the right. There is a corresponding set of antiparticles, quarks and leptons carrying the opposite charge. The weak interactions between quarks of the same generation are highly favored. For example, a t quark preferentially decays to a b quark, a key to the ultimate identification of the t quark. In the standard model of the electroweak interactions, neutral currents cannot cause a change in the quark flavor. *See* ANTIMATTER; ELECTROWEAK INTERACTION; ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; LEPTON; QUARKS; STANDARD MODEL; WEAK NUCLEAR INTERACTIONS.

Val L. Fitch

Flax

The flax plant (*Linum usitatissimum*) is the source of two products: flaxseed for linseed oil and fiber for linen products. Plants with two distinct types of growth are used for seed and fiber production (**Fig. 1**, but fiber is sometimes a by-product of linseed production, and seed is an important by-product of fiber flax. Flax is an annual plant with a stem size that varies between about $^1/_{16}$ and $^1/_4$ in. (1.6 and 6.4 mm) in diameter. The flax flower has five petals, which are blue, white, or pale pink (**Fig. 2**). The color of the flower depends on the variety. Seeds of most varieties are varying shades of brown and yellow. Flax must be grown in rotation with other crops and not planted on the same land more frequently than once every 4 to 6 years; otherwise disease-producing organisms build up in the soil and yields may be greatly reduced unless resistant varieties are grown. *See* SEED; STEM.

Flax for fiber. Flax for fiber requires fertile, well-drained, and well-prepared soil and a cool, humid climate. It is planted in the early spring at 80 lb/acre (70 kg/hectare) or more. The high rate of seeding provides a dense stand of slender stems that branch only at the very top. Some chemical weed control may be necessary. A few weeds such as wild oats, cereal crops, perennial thistles, or morning glories will destroy the value of the crop for fiber, and all fields badly infested with weed seeds of any kind must be avoided for planting fiber flax. *See* AGRICULTURAL SOIL AND CROP PRACTICES; HERBICIDE.

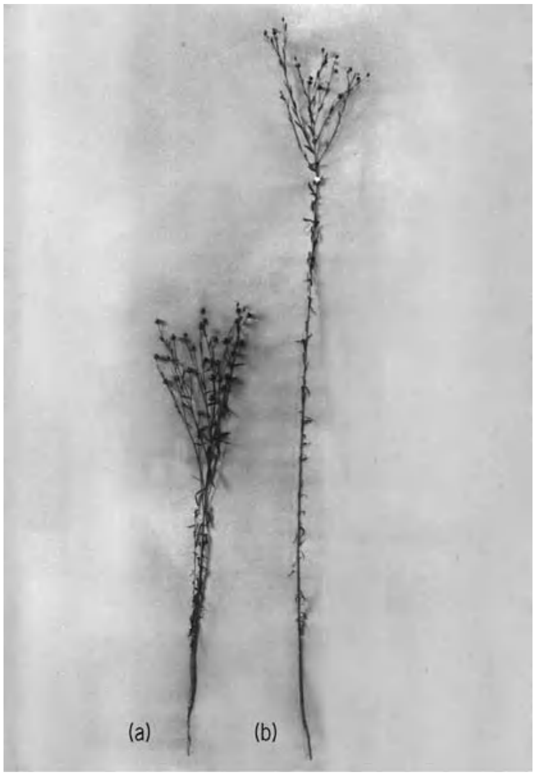


Fig. 1. Types of flax plant. (*a*) Seedflax. (*b*) Fiber flax. (*USDA*)

Fig. 2. Flowering top of flax plant showing buds and seed capsules. (*USDA*)

Harvesting. When about half the seed pods (bolls) have turned yellow, harvesting must be started. If the bolls are allowed to reach full maturity before harvesting, quality is greatly reduced.

Fiber flax is harvested by a machine that pulls it out of the ground and ties it in bundles. Some flax is still pulled by hand. Stems are kept parallel while the seed is removed; bundles of the straw are then packed vertically in concrete tanks and covered with warm water to ret, a process in which bacterial action loosens the fiber in the stem. The bundles of wet, retted straw are set up with the tops held together and the root ends spread out so the bundles form “gaits” or “wigwams” to dry in the sun. The dried flax is retied in bundles and stored under cover until it is scutched, a process in which the straw is first crushed to break the woody portion of the stem into small pieces (called shives) and is then subjected to a beating operation that removes the outer bark and the woody parts from the line (long parallel) fiber. During the beating operation, short tangled fiber called tow also comes out of the line fiber with the shives and must be further cleaned. In Western Europe the scutching operation is mechanized. There, the straw is spread in a thin layer, held near the center by a gripping device on a narrow belt, and carried between fluted rollers to break the woody parts into shives. The shives are then beat out of one end, the device grips the cleaned portion, and the other end is cleaned, leaving long (16–40 in. or 40–100 cm) strands. *See* AGRICULTURAL MACHINERY.

Uses. The best-known use for flax fiber is in the manufacture of fine linen fabrics; other uses are linen thread, linen twine, toweling, and canvas. Seed from fiber flax is used for replanting and for oil. The major producer of flax fiber is Russia, but the world's best fiber comes from Belgium and adjoining countries. Most Irish linen is manufactured from fiber produced in Belgium. *See* LINEN; NATURAL FIBER; TEXTILE.

Flax for seed. The sowing rate for seedflax is lower than for fiber flax, about 40–50 lb/acre (45–55 kg/ha). The varieties used produce short straw and a relatively high yield of seed with 35–44% oil. Machines used for planting and harvesting cereals are also used for flax for seed. Planting is usually done in the early spring on a well-prepared seedbed on land relatively free from weed seeds. A small acreage of flax is sown in the fall in Texas. Herbicides may be used to control weeds. *See* CEREAL; FAT AND OIL.

The world's major producers of flaxseed are Argentina, Canada, India, Russia, and the United States. Most of the flaxseed acreage in the United States is in Minnesota, North Dakota, and South Dakota.

The principal uses of linseed oil are in the manufacture of protective coatings—paint, varnish, and lacquer. It is also used in linoleum, printer's ink, and patent and imitation leathers, and as core oil for making sand forms in metal casting. Uses that promise to become of great importance are as antispalling compounds for concrete highway construction and as curing agents for concrete. The linseed cake remaining after the oil is extracted is used as a high-protein livestock feed. Large quantities of clean (weed-free) straw from seedflax are used in the production of a high-grade paper. This product is used to make cigarette and other fine papers. Elton G. Nelson; Joseph O. Colbertson

Diseases. Of nearly two dozen diseases of oil and fiber flaxes throughout the world, fungi cause about 16 diseases, a virus causes crinkle, a mycoplasma-like organism causes aster yellows, a chlorophylless seed plant known as dodder (*Cuscuta epilinum*) is parasitic on flax stems, and several nonparasitic agents account for other diseases such as heat canker and iron chlorosis.

Diseases of flax can result either in complete losses of stands or in reduced yields at harvest, depending upon the disease, cultural practice, seed quality, environment, and resistance in the flax cultivar. Diseases destructive in one locality may be of negligible importance in another. For example, pasmo is important in Minnesota and the Dakotas, but it is of little importance in California and is unknown in Mexico; similarly, stem break (caused by *Polyspora lini*) is troublesome in Europe but the disease is rarely found in the United States.

Rust (*Melampsora lini*) is potentially the most destructive disease of flax wherever flax is grown in the world, primarily because its airborne spores can generate an epidemic rapidly within a season in one area. Severe rust occurred in North America in 1939–1941, 1944–1945, 1949–1951, and 1963, following the introduction of cultivars that proved to be susceptible to rust. Because no one cultivar is resistant to all of the more than 300 races of rust, cultivars must be grown that are resistant to the races prevalent in a given area, and this is the only practical method of control.

Wilt caused by *Fusarium oxysporum* was the limiting factor in flax production in the United States early in this century. Complete losses in seedling stands occurred in commercial cultivars (**Fig. 3**). As wilt resistance was bred into cultivars, some with moderate seedling resistance wilted after flowering, resulting in 8–20% losses in yield. The fungus is seed- and soil-borne, and once established in soil it survives there for decades, even as long as 50 years in the absence of flax. This makes crop rotation and

Fig. 3. Wilt resistance plot in Minnesota. Three rows of susceptible cultivar in the center are compared with rows of resistant cultivars on each side, in a plot where flax has been grown each year for 50 years.

Fig. 4. Pasmo disease of flax. Dark-colored bands indicate places of infection with the pasmo fungus which girdles stems and kills plants.

Fig. 5. Blight of flaxseed caused by several fungi. (*Minnesota Agricultural Experiment Station*)

fungicides ineffective; breeding for resistance is the only feasible method of control, and this method is eminently successful today.

Pasmo, caused by *Septoria linicola*, occurs in Europe, North and South America, and New Zealand, and was introduced into the United States from Argentina about 1915. It became destructive within 30 years. The stems turn brown, often in bands (**Fig. 4**), and leaves fall off, and at harvest the seeds are shriveled. Because no resistant cultivars are known, crop rotation and sanitation are the only possible control measures.

Seedling blight caused by several fungi (*Alternaria* spp. *Colletotrichum lini, Pythium* spp., and *Rhizoctonia solani*) and root rot caused by *R. solani* can be important at certain times and places. Sometimes seed condition determines whether seedling blight and root rot occur. Damaged flaxseed, known also as cracked, chipped, split, fractured, weathered, or blighted seed, is common in all flax-growing regions of North America (**Fig. 5**). Sowing damaged seed can result in 10–50% lower stands and sometimes 20% lower yields because organisms in the soil can enter openings in damaged seeds or infect crippled seedlings growing from damaged seeds. The best control for seedling diseases is to select sound seeds and treat them with an appropriate fungicide. Some seed-borne, seedling blight fungi are also controlled by fungicides.

Only one virus is important on flax, the oat blue dwarf virus, which causes crinkle. A mycoplasma-like organism (*Spiroplasma kunkelii*) causes aster yellows in flax. Both of these pathogenic agents reduce seed yields and are transmitted by the same vector, the aster leaf-hopper. Control is by growing resistant cultivars. *See* PLANT VIRUSES AND VIROIDS.

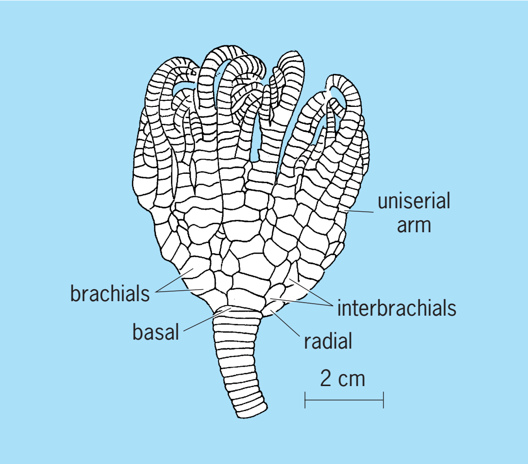
There are no important bacterial or nematode diseases of flax. Flax is susceptible to both hail and insect damage which predispose plants to disease. *See* PLANT PATHOLOGY.

Thor Kommedahl

Bibliography. H. H. Flor and V. E. Comstock, Flax cultivars with multiple rust-conditioning genes, *Crop Sci.,* 11:64–66, 1971; T. Kommedahl, J. J. Christensen, and R. A. Frederiksen, *A Half Century of Research in Minnesota on Flax Wilt Caused by Fusarium oxysporum,* Minn. Agr. Exp. Sta. Tech. Bull. 273, 1970; P. Neergaard, *Seed Pathology*, vols. 1 and 2, 1977.

Flexibilia

An extinct subclass of stalked or creeping Crinoidea which includes some 50 Paleozoic genera, ranging from Ordovician to Permian times. A flexible tegmen was present with open ambulacral grooves. One of the five oral plates served as a madreporite. The anus was at the tip of a short siphon. About three of the lower brachial ossicles of each arm were loosely incorporated into the dicyclic calyx. The three infrabasals were sometimes obscured by the stem, but the five basals and five radials remained conspicuous and movably articulated. The uniserial arms lacked pinnules but branched freely, arching inward to form a globular crown (see **illus.**). The cylindrical stem lacked cirri and was sometimes discarded, as in *Edriocrinus*, to produce a creeping adult. *See* CRINOIDEA; ECHINODERMATA.



Talanterocrinus sp., a member of the Flexibilia.

Howard B. Fell

Flexible manufacturing system

A factory or part of a factory made up of programmable machines and devices that can communicate with one another. Materials such as parts, pallets, and tools are transported automatically within the flexible manufacturing system and sometimes to and from it. Some form of computer-based, unified control allows complete or partial automatic operation.

Programming. The programmable machines and devices are numerically controlled machine tools, robots, measuring machines, and materials-handling equipment. Each programmable machine or device typically has its own controller which is, in effect, a dedicated digital computer; programs must be written for these controllers. For example, a numerically controlled machine needs a program which supplies the sequence of machining operations to be carried out on a given part. A robot needs a program which specifies where and how it must move and employ its gripper. These programs are usually written in special-purpose languages designed to handle the geometry and machining. Increasingly, numerically controlled machines are being programmed by graphical presentations on computer screens, that is, graphical computer interfaces. This allows the programmer to follow the machining operation and specify desired operations without the need for statements in a programming language. Robots have usually been programmed by so-called teaching, where the robot is physically led through a sequence of movements and operations; the robot remembers them and carries them out when requested. Robots are also programmed with robot-programming languages and, increasingly, with graphical computer interfaces. Similar choices in programming method are available for other devices. Programming a flexible manufacturing system may involve several different languages and methods, resulting in complicated and expensive software development and maintenance. *See* COMPUTER GRAPHICS; COMPUTER NUMERICAL CONTROL; COMPUTER PROGRAMMING; INTELLIGENT MACHINE; MATERIALS-HANDLING EQUIPMENT; PROGRAMMING LANGUAGES; ROBOTICS.

Communication. The programmable machines and devices communicate with one another via an electronic connection between their controllers. Increasingly, this connection is by means of local-area networks, that is, communication networks that facilitate high-speed, reliable communication throughout the entire factory. These connections allow the programs in the various programmable devices to interact. For example, when a numerically controlled machine has completed a part, it can ask the robot to put it on a vehicle. *See* LOCAL-AREA NETWORKS.

Material transport system. The automatic material transport system is usually a guided, computer-controlled vehicle system. The vehicles are usually confined to a fixed network of paths, but typically any vehicle can be made to go from any point in the network to any other point. The network is different from a classical assembly line in that it is more complex and the flow through it is not in one direction. Paths in the system can be tracks to which the vehicles are confined or a marker, such as wire buried in the floor, that a vehicle senses and follows.

Since the vehicles must transport various kinds, shapes, and sizes of objects to various places, a way to accommodate this automatically is needed. One widely used method is to employ pallets standardized in the sense that common fittings allow each pallet to mount onto each vehicle and each location. For example, all pallets might be designed to mount on four vertical conical pins, or all might have the same slots machined into their sides. There are other possibilities, but in all cases a standardized mounting is the key. Beyond this, it must be possible to attach each object to be transported to a pallet, sometimes automatically and sometimes manually. Typically a part is attached to a pallet that is loaded into the system, picked up by a vehicle, and loaded into the numerically controlled machine.

Unified control. Commands and orders to the flexible manufacturing system are sent to its computer-based, unified control. For example, an order can be given to make 25 parts of type 76-998. The control, in turn, issues orders for the transport of various kinds of material, the transfer of needed programs, the starting and stopping of programs, the scheduling of these activities, and other activities. Typically, the control is in a central computer dedicated mainly to it; this computer is usually not a device controller.

Flexibility. Flexible manufacturing systems are flexible in the sense that their device controllers and central control computer can be reprogrammed to make new parts or old parts in new ways. They can also often make a number of different types of parts at the same time. However, this flexibility is limited. Each system is designed with a certain family of parts in mind, for example, axles. It can often be reprogrammed to a make a new kind of axle, and it can usually make a number of different kinds of axles at the same time. But its flexibility would almost certainly not extend to making engine blocks.

Advantages. Various rationales for flexible manufacturing systems have been given. The reduction of the work force involved is rarely the most important one. In fact, direct labor costs are a relatively small part of total costs in many manufacturing systems, particularly those that are candidates for conversion to flexible system. Thus, direct labor reduction usually has a minor impact on total costs.

Because the flexible manufacturing system is under continuous computer control, the movement of parts, material, and tools is efficient in the sense that neither time nor space is wasted. If a part needs to be moved it is done immediately. Such a performance is difficult to obtain in nonautomated systems other than highly structured, inflexible ones. Therefore, the number of partially completed parts that have to be in the system (referred to as work in progress) for a given level of production is significantly reduced, resulting in major cost reductions. For example, if producing 1000 axles a month requires a work in progress of 10,000 partially completed axles and if introduction of a flexible manufacturing system reduces work in progress to 2000 axles, the cost of carrying an inventory of 8000 partially completed axles will be saved. *See* INVENTORY CONTROL.

Similarly, the processing time from raw material to finished part is reduced, and this allows for a reduced inventory of finished products and quicker response to market demands. The first reduces inventory costs, and the second, though more difficult to quantify, improves competitiveness.

Computer control also reduces the idle time for the machines. Moreover, flexible manufacturing systems are increasingly designed to continue unattended during nights, weekends, and holidays. This is an important source of additional production. Fewer machines and less floor space are needed.

Flexible manufacturing systems typically produce a better product for several reasons. The controlling computers execute the required programs day and night with few mistakes. Since the mechanical handling of parts is automatic, it must be designed to be done repeatedly with very few mistakes. Typically, quality control measurements are built into the system. This is important because it allows deviations to be detected, diagnosed, and corrected at once, often with little interruption in production. Moreover, computer control allows problems to be anticipated and preventive actions to be carried out. For example, many systems keep a history for each tool and can replace it in anticipation of breakage. In addition to a better-quality product, this reduces the number of parts that have to be discarded or reworked, and translates into higher production or a smaller system. *See* QUALITY CONTROL.

In addition, flexible manufacturing systems allow production levels to be easily and quickly changed; the production of sample new designs can be intermingled with ongoing production, and product composition can be readily changed. Product changeovers are also easier. Furthermore, they allow a reduction in indirect labor because complete information about manufacturing operations is available immediately from the computer.

Disadvantages. Even though flexible manufacturing systems have existed for over 40 years, they still are complex systems that are difficult to create, especially if flexibility is emphasized. For example, they require high skills and training, particularly with respect to computers. Development, maintenance, and modification of all needed computer programs often present major difficulties, which is the main reason that flexibility of programmable devices is often not fully exploited.

Flexible manufacturing systems are best suited for specialized manufacturing. On the one hand, they are not viewed as a replacement for so-called hard automation as exemplified by the assembly line. On the other hand, they are not intended to compete with the small-job shop. In the first case, flexible manufacturing systems have unneeded and expensive flexibility, and in the other they are not flexible enough.

It has also been widely noted that introduction of flexible automation usually requires a major change in the way that a firm organizes and manages its staff; the difficulty of defining and making these changes is comparable with that presented by programming.

Renewed attention in the late 1980s to the management of manufacturing systems with automation levels significantly lower than those of flexible manufacturing systems, with particular attention to personnel management, resulted in important improvements in productivity and quality.

Computer-integrated manufacturing. Flexible manufacturing systems are part of a larger computer-based technology, termed computer-integrated manufacturing (CIM), which encompasses more than the movement and processing of parts on the factory floor. For example, computer-aided design (CAD) of products is a major part of CIM, as are production and capacity planning. Indeed, the scope of CIM is wide, including the computerization of a manufacturing company and its suppliers and customers. *See* AUTOMATION; COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-INTEGRATED MANUFACTURING.

Arch W. Naylor

Bibliography. D. M. Considine and G. D. Considine (eds.), *Standard Handbook of Industrial Automation*, 1986; B. Luggen, *Flexible Manufacturing Cells and Systems*, 1990; R. Maleki, *Flexible Manufacturing Systems: Technology and Management*, 1990; A. Ross, *Dynamic Factory Automation*, 1992; H. Tempelmeier and H. Kuhn, *Flexible Manufacturing Systems: Decision Support for Design and Operation*, 1993; D. Tijunelis and K. E. McKee, *Manufacturing High Technology Handbook*, 1987; H. J. Warnecke and R. Steinhilper (eds.), *Flexible Manufacturing Systems*, 1989.

Flight

The movement through the atmosphere or space, sustained by aerodynamic reaction or other forces. Animal flight includes gliding and flapping flight. Four animal groups evolved flapping flight, namely insects, pterosaurs, birds, and bats. Flapping flight in vertebrates was probably preceded by gliding; in insects it may have originated by leaping and gliding, by surface skimming on water, or (if small enough) by passive floating in the air. Flying insects show greater variation than flying vertebrates, and their flight spans a wider range of Reynolds numbers, which is the ratio of inertial forces to viscous forces in the flow. Flight of tiny insects is in the lower range of Reynolds numbers, where viscous forces are dominant, whereas large insects and vertebrates operate in the higher range, where inertial forces are important.

Flight is very expensive in terms of energy cost per unit time. However, flying animals can travel several times faster than nonflying ones, and the cost of carrying a unit of weight of an animal through a unit distance (cost of transport) is lower for flight than for running; but higher than for swimming.

Characteristics. The flight characteristics of large insects and vertebrates can be understood by aircraft aerodynamics. In steady level flight, an animal and an aircraft must generate forces that support weight against gravity and provide propulsive thrust against drag forces. The forces acting on an airfoil (the shape of the cross section of the wings) moving through the air depend upon the flow pattern around it. Because of the asymmetric profile on an airfoil, the air flowing over the upper surface travels farther and flows faster than air passing underneath. According to Bernoulli's principle, pressure falls when speed rises in a moving fluid, resulting in a pressure difference between the upper and lower sides of the airfoil **(Fig. 1*a*)**. This pressure difference has to disappear gradually toward the wingtips, and some air will flow upward around the wingtips. The air moves downward behind the wing as trailing vortices (Fig. 1*b*), and the reaction of this momentum flow is experienced by the wing as lift. The stronger the vortex, the greater the lift generated, but with some energy loss to drag. The lift force is responsible for weight support (its vertical component) and thrust or drag (its horizontal component). *See* AERODYNAMIC FORCE; AIRFOIL; BERNOULLI'S THEOREM.

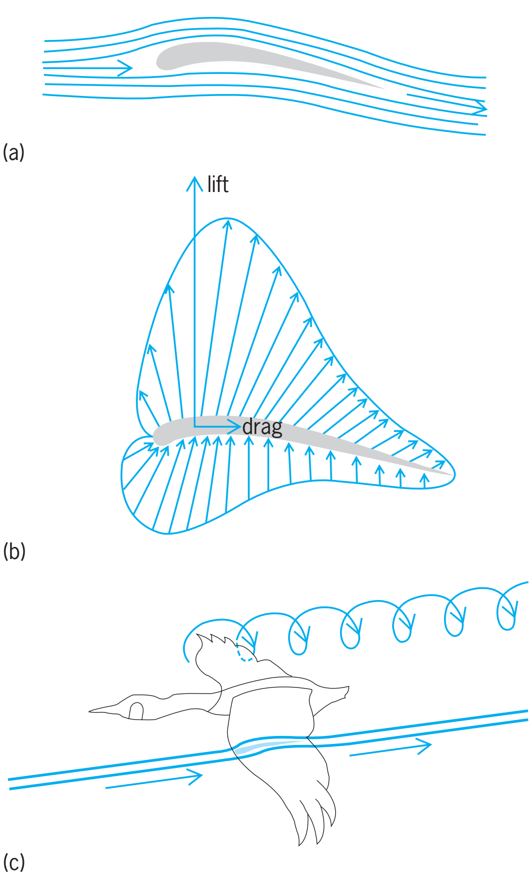


Fig. 1. Aerodynamics of flight. (*a*) Airflow around a typical wing profile. (*b*) Pressure distribution around a typical wing profile. (*c*) The difference in pressure disappears toward the wingtips as trailing vortices in the wake. (*After M. Brooke and T. Birkhead, eds., Cambridge Encyclopedia of Ornithology, Cambridge University Press, 1991*)

Energy expenditure. Energy must be expended to generate the trailing vortices in the wake and to overcome friction on the wing and body surfaces. These energy losses are experienced as drag forces, which act parallel to the direction of movement of the airfoil. The drag caused by trailing wingtip vortices is conventionally termed induced drag, the frictional drag of the wings is termed profile drag, and the frictional drag of the body is the parasite drag. The magnitude of lift and drag varies with the velocity of the flow, the size and shape of the wing, and the angle at which the flow meets the airfoil (angle of incidence). When the stalling angle is exceeded, the airstream separates from the upper surface of the wing with a sudden fall of lift and increase of drag. A good airfoil maximizes the pressure difference between its upper and lower sides and minimizes the drag: the lift:drag ratio of avian wings averages between 5:1 and 10:1.

Since drag is a retarding force, the animal must either descend (glide) through the air at such an angle that a component of its weight balances the drag, or do mechanical work with its flight muscles (flap its wings) to overcome it. The rate at which this work is done is the mechanical power required to fly, and it equals speed times drag [measured in watts (W)]. The flight muscles also produce heat when they contract, so the total metabolic power expenditure is the sum of this heat loss and the mechanical power. The metabolic power is estimated to be four to five times the mechanical power, and is dependent on the size of the animal. *See* WORK.

Gliding flight. Compared with active flight, gliding flight is very inexpensive, and is found in a wide range of animals, such as squirrels, marsupials, lizards, frogs, fishes, and even one snake. It is the main component in soaring flight used by many birds and some bats, when the animals use thermals or updrafts. Gliding in birds costs only two to three times the basal metabolic rate, because the flight muscles do not perform any mechanical work other than for stability and control of movements, and produce mostly static forces to keep the wings down on the horizontal plane, opposing the aerodynamic force.

When gliding, the wings of the animal leave behind a continuous vortex sheet that rolls up into a pair of vortex tubes (wingtip vortices), as in fixed-wing airplanes (Fig. 1*b*). The lift produced balances the animal's weight, but potential energy must be used to overcome the drag and the animal loses height. An animal gliding at steady speed descends at an angle to the horizontal (glide angle), established by the ratio of lift to drag (glide ratio). The best glide ratios in birds range from 10:1 to 15:1 for vultures and birds of prey and reach 23:1 in the wandering albatross, whereas modern gliders can achieve 45:1. An animal cannot glide more slowly than its stalling speed, which in birds can be reduced by splaying the wingtip primaries, or by spreading the alula (a digit of the wing) at the wrist, or both. An animal can increase its gliding speed by flexing the wings and reducing the wing area.

Flapping flight. In horizontal forward flight, the velocity of the air that meets a wing section depends on the forward speed of the animal and the flapping speed of the wing section. The local forces on the wings are therefore different along the wing and vary throughout the wing-beat cycle. In the downstroke the inner parts of the wing produce small lift and drag forces, and the outer parts mainly lift and thrust. The structure of the wake varies with flight speed. As the wings travel in flapping flight, they leave behind them a vortex sheet that is formed during the first half of the lift-producing downstroke. During the upstroke the wings are more or less flexed to reduce drag, have no bound vorticity, and contribute little or no lift. The rapid acceleration and deceleration of the wings at the beginning and end of the downstroke alters the circulation on the wing, and the shed trailing vortices then close into vortex loops, one for each downstroke (**Fig. 2***a*). During faster flight at low wing-beat amplitude and frequency, the vortex wake consists of a pair of continuous undulating vortex tubes or line vortices behind the wingtips (Fig. 2*b*), as in kestrels in cruising flight, indicating that also the upstroke provides useful aerodynamic forces.

Hovering flight represents the most expensive form of animal flapping flight. The essence of hovering is to produce a vertical force to balance body weight. The wake consists of a chain of vortex rings continuously shed during the wing strokes. In hummingbirds and insects, lift is produced during both the downstroke and upstroke of the wings (symmetrical hovering), and two vortex rings are produced during each wing stroke (Fig. 2*c*). In other hovering animals the wings are flexed during the upstroke (asymmetrical hovering), and the rings are produced during the downstrokes only (Fig. 2*d*).

Wing shape and flight behavior. Different foraging behavior and associated flight modes put different demands on body size and wing form. The relative size and shape of the wings can be quantified by the wing loading and the aspect ratio. Wing loading is estimated as the ratio of body weight to wing area [in units of newtons (N) per square meter]. It increases with body size and ranges from around 10 N/m^2 in small flycatchers and owls to 230 N/m^2 in the whooper swan. Flight speed is proportional to the square root of wing loading, and the radius of a banked turn is proportional to wing loading. Slow and maneuverable flight is thus obtained by a low wing loading.

Aspect ratio is calculated as the wing span squared divided by wing area. It is dimensionless and ranges from around 4.5 in the great tit and some galliforms and tinamous to around 18 in albatrosses. Aspect ratio is a measure of the aerodynamic efficiency, and lower energy cost of flight is obtained in animals with wings of high aspect ratio, and particularly in combination with a low wing loading (long, narrow wings).

Many seabirds, swifts, and swallows have long and narrow wings, and they can thus afford to fly continuously during foraging. Animals with short and narrow wings, that is, with higher wing loadings (including loons, mergansers, ducks, auks, and free-tailed bats) are capable of fast and, for their speed, rather economical flight, used for commuting and migration.

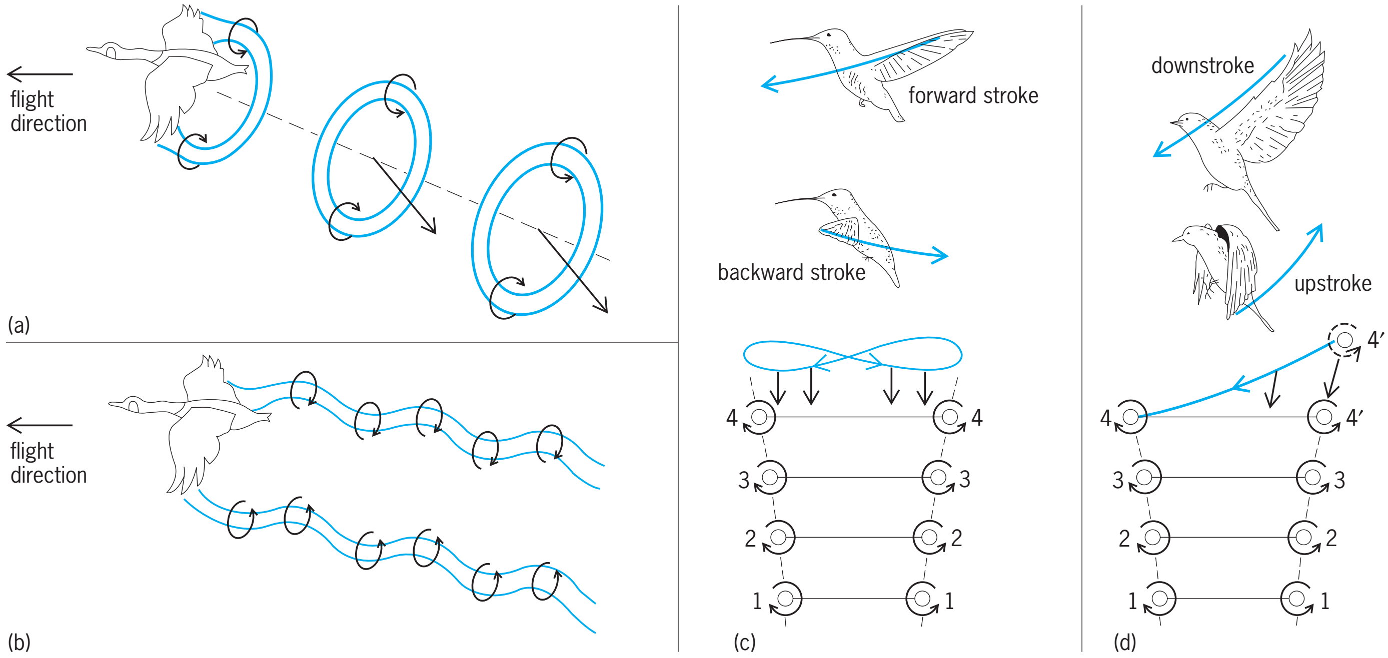


Fig. 2. Vortex wakes behind a bird in flapping flight, and below a hovering bird. (*a*) In slow flight, vorticity is shed only during the downstroke, resulting in a stack of vortex rings. (*b*) In faster flight, lift can be generated also during the upstroke, and the wake consists of an undulating pair of line vortices (*after M. Brooke and T. Birkhead, eds., Cambridge Encyclopedia of Ornithology, Cambridge University Press, 1991*). (*c*) Formation of vortex rings in four wingstrokes in symmetrical hovering and (*d*) in asymmetrical hovering. (*After U. Norberg, Vertebrate Flight, Springer, 1990*).

Soaring birds and animals flying among vegetation benefit from low weight and broad wings for slow flight and high maneuverability. Birds in this group often have a large alula and the ability to separate the wingtip primaries, so as to increase lift at large angles of incidence, thus permitting slow speeds. The very low aspect ratios in animals flying among vegetation are connected with high flight costs, and the animals are, therefore, usually perchers or spend much of their foraging time walking, climbing, and hanging. *See* AERODYNAMICS; AVES. Ulla M. Norberg

Bibliography. M. Brooke and T. Birkhead (eds.), *Cambridge Encyclopedia of Ornithology*, 1991; R. F. Johnson (ed.), *Current Ornithology*, vol. 5, 1988; U. M. Norberg, *Vertebrate Flight*, 1990; C. J. Pennycuick, *Bird Flight Performance*, 1989.

Flight characteristics

The behavior of an aircraft while airborne. An aircraft, while in the air, is free to move in six degrees of freedom (**Fig. 1**). The aircraft can translate along any of its three axes, X, Y, and Z. The corresponding translational velocities are identified as U (forward velocity), V (side velocity), and W (up or down velocity), respectively. In addition, the aircraft can rotate about any of its three axes, X, Y, and Z. The corresponding rotational velocities are labeled as P (roll rate), Q (pitch rate), and R (yaw rate) respectively. *See* DEGREE OF FREEDOM (MECHANICS).

The mission of most aircraft is to carry a payload from one point on Earth to another. To carry out that mission the aircraft must satisfy several requirements: (1) It should be able to develop enough lift to overcome its weight. (2) It should be able to develop enough thrust to overcome its drag. (3) It should have sufficient fuel on board to cover a useful range with a useful payload. (4) It should be able to take off and land on reasonably sized airfields. (5) It should be easily and safely maneuverable from one flight condition to another.

For any given airplane, these five requirements are given specific numerical values which depend on the mission of the aircraft. To predict whether or not a given aircraft design can meet these requirements, its flight characteristics are separated into two categories: performance and flying-quality (also called handling-quality) characteristics. Both are subject to airworthiness regulations by civilian and military authorities because of safety concerns. These regulations form important constraints on the design of aircraft and aircraft systems. *See* AIRCRAFT DESIGN.

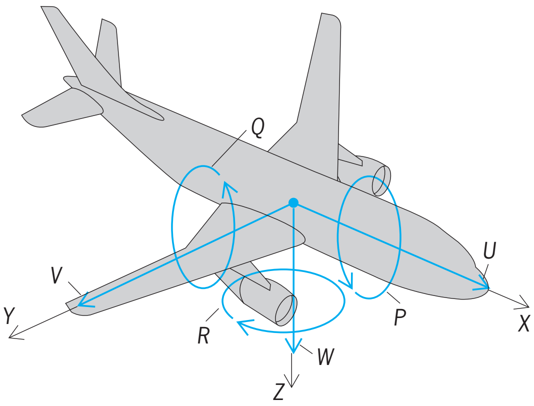


Fig. 1. Aircraft axes *X*, *Y*, *Z* and degrees of freedom *U*, *V*, *W*, *P*, *Q*, *R*.

Performance characteristics. To predict and describe aircraft performance characteristics, it is usually sufficient to address the following performance categories: takeoff distance, climb rate or climb gradient, time to climb to altitude, maneuverability, cruise speed and altitude, payload-range and fuel used, loiter, descent rate, descent time, and landing distance. These performance categories are normally summarized with the help of a flight envelope (**Fig. 2**) and a mission profile (**Fig. 3**). Cruise speed and payload-range will be discussed in some detail to highlight important aircraft design principles and design parameters.

Cruise speed. The cruise speed of an airplane depends on weight, drag, thrust, and altitude. The cruise drag, which must be overcome by thrust, is itself the sum of two contributions: drag at zero lift (mostly caused by friction of the air over the surfaces of the airplane) and drag due to lift (also called induced drag). The thrust required to overcome the cruise drag (in pounds or newtons) is referred to as the required thrust (T_{required}). The combined engine thrust should be able to overcome the cruise drag. The combined engine thrust at some thrust setting and altitude is referred to as the available thrust ($T_{\text{available}}$).

At the cruise speed, the thrust required and the thrust available are equal to each other. The cruise speed U can be determined from the airplane cruise weight; the zero-lift drag coefficient of the airplane; the air density at cruise altitude; the dimensions of the wing; the wing aspect ratio; and Oswald's efficiency factor, which accounts for the fact that lift distributions on practical wings are not perfectly elliptical. The cruise speed U can be determined graphically from a plot of required and available thrust (**Fig. 4**). *See* AERODYNAMIC FORCE; AIRFOIL; PROPULSION; THRUST.

Payload-range. Most aircraft are designed to carry a specified payload over a given distance, at a given speed and altitude, while using a minimum amount of fuel. It is important to understand the relationship between various weights [in pounds-force or newtons]. In the equation below W_{TO} is the airplane

$$W_{TO} = W_{OE} + W_{PL} + W_F$$

takeoff weight, W_{OE} is the airplane operating empty weight, W_{PL} is the payload weight, and W_F is the mission fuel weight. All airplanes are designed to some maximum allowable takeoff weight mostly to prevent overloading. The operating empty weight normally consists of the actual empty weight plus the weight of the crew needed to operate the airplane. The fuel weight consists of the fuel used during the

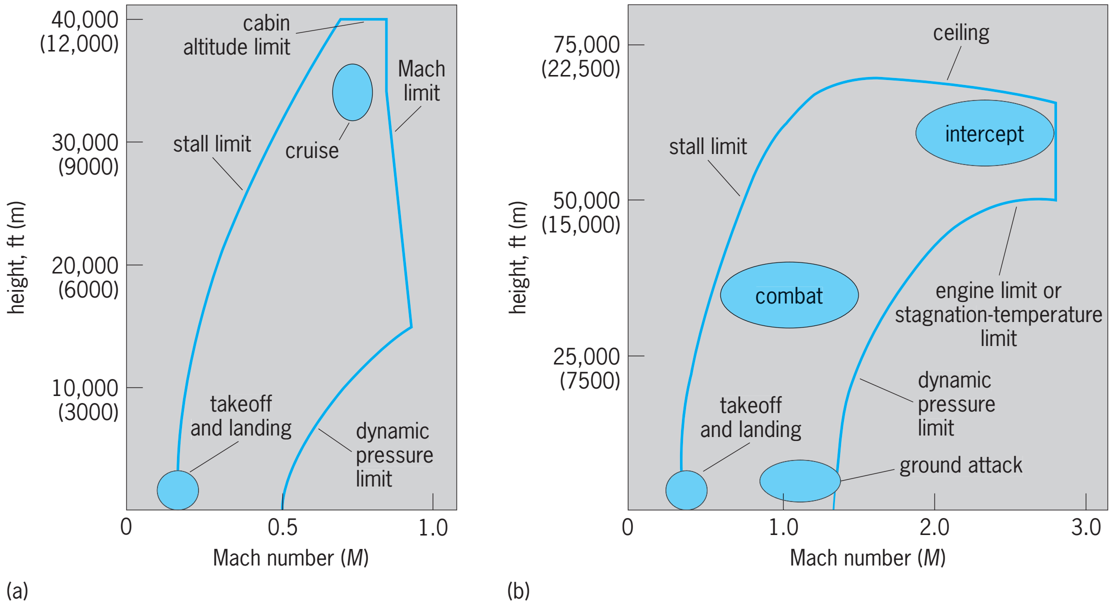


Fig. 2. **Examples of operational flight envelopes. (*a*) Civilian airplane. (*b*) Military airplane.**

mission plus the reserve fuel required by civil or military regulations.

The maximum allowable takeoff weight and the empty weight are defined in the flight manual for any specific airplane. The difference between the takeoff weight and the operating empty weight is referred to as the useful weight. From Eq. (1) it follows that the useful weight is the sum of payload weight and mission fuel weight. Operators of airplanes must decide which combination of payload and fuel to carry for a given mission. It is usually not acceptable to exceed the maximum allowable takeoff weight. Most airplanes also have a limited amount of fuel volume available.

The range (in nautical miles or kilometers) which can be covered by a jet airplane while cruising at constant speed and angle of attack may be computed from the airplane cruise speed, the specific fuel consumption, the lift-to-drag ratio in cruise, and the ratio of airplane weight at the beginning of cruise to that at the end of cruise. Roughly, the weight at the beginning of cruise may be thought of as equal to the takeoff weight. The weight at the end of cruise is then lower by the amount of fuel actually used. *See* SPECIFIC FUEL CONSUMPTION.

High cruise speed and low fuel consumption, coupled with a high lift-to-drag ratio and a high weight ratio, are conducive to long range. Operators of commercial and military transports are interested in the relationship between range and payload. This relationship is normally represented in the form of a payload-range diagram. The maximum range with full payload is referred to as the harmonic range. The range with maximum fuel (from a volumetric viewpoint) can be attained at a reduced payload only. The maximum range at zero payload is called the ferry range.

Flying-quality characteristics. In the case of airplanes that are flown by human pilots, the interaction between pilot cockpit control inputs and airplane response to the cockpit control inputs must be such that the pilot can achieve the mission objectives with reasonable physical and mental efforts. In other words, the airplane is required to have acceptable flying qualities (sometimes also called handling qualities) anywhere inside the operational flight envelope.

The flying qualities of an airplane must be such that

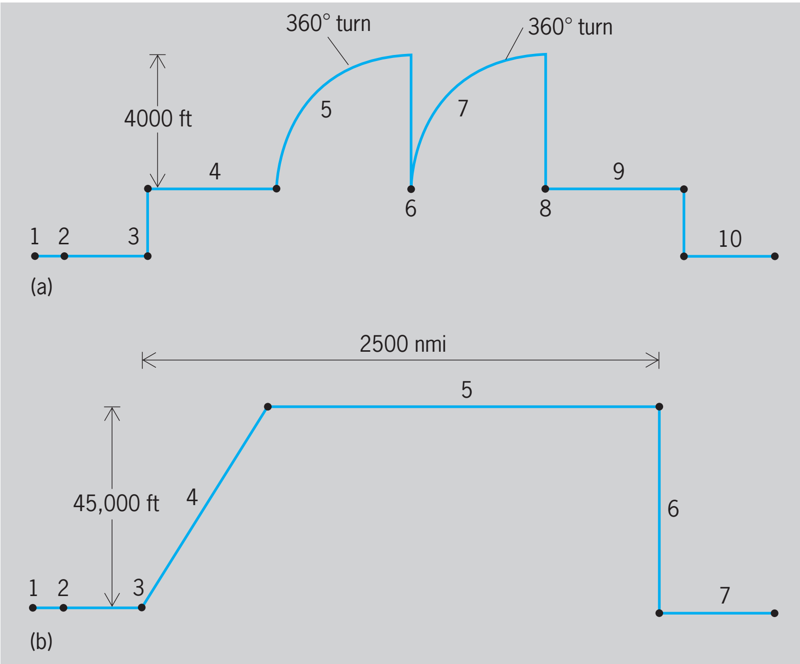


Fig. 3. **Examples of typical mission profiles and flight phases. (*a*) Military attack airplane. Numbers indicate flight phases: (1) engine start and warmup; (2) taxi; (3) take off and accelerate to 350 knots (650 km/h) at sea level; (4) dash 200 nm (370 km) at 350 knots; (5) 360°, sustained, 4.5-*g* turn, including a 4000-ft (1200-m) altitude gain; (6) release two bombs and fire 50% ammunition; (7) 360°, sustained, 4.5-*g* turn, including a 4000-ft altitude gain; (8) release two bombs and fire 50% ammunition; (9) dash 200 nm at 350 knots; (10) land, taxi, shut down. (*b*) Civilian, passenger-transport airplane. Numbers indicate flight phases: (1) engine start and warmup; (2) taxi; (3) take off; (4) climb to 45,000 ft (13,500 m); (5) cruise; (6) descend; (7) land, taxi, shut down. 2500 nm = 4630 km.**

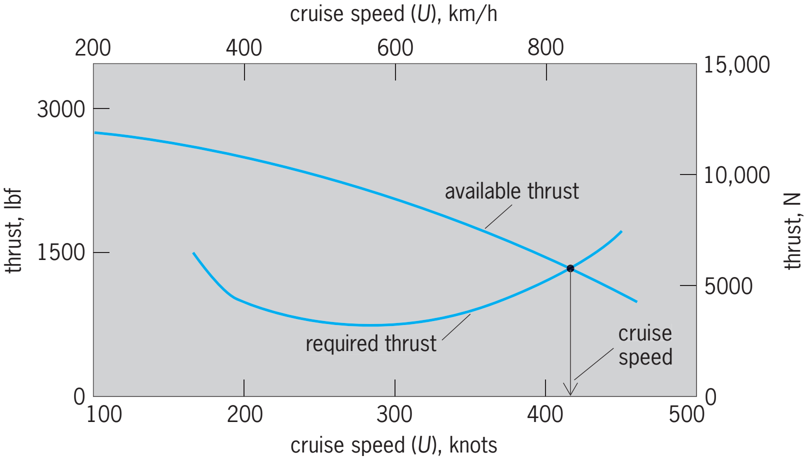


Fig. 4. Determination of cruise speed for a typical jet airplane. Thrust setting is for maximum continuous thrust in cruise. Data apply to one weight only.

the following characteristics are present anywhere inside the operational flight envelope:

1. The airplane must have sufficient control power to maintain steady-state, straight-line flight (level or not) as well as steady-state maneuvering flight consistent with mission objectives.

2. The airplane must be maneuverable from one (steady-state) flight condition to another.

3. The airplane must have sufficient control power to accomplish the transition from ground operations to airborne operations (takeoff run, lift-off, and initial steady-state climb), and the transition from airborne operations to ground operations (steady-state approach, touch-down, and landing run).

These three characteristics must be present whether the engine or engines are operating or not. Included are certain conditions of asymmetrical power which can be expected to occur. The whole is referred to as the minimum control speed condition.

The physical pilot efforts required to fly an airplane while meeting its mission objectives are expressed in terms of cockpit control activity and cockpit control forces. The cockpit control force magnitudes and their variation with certain motion parameters (examples are speed, load factor, and sideslip) are important to pilots. Examples of flying-quality parameters that deal with physical pilot efforts are maximum required stick force, stick force per g, stick-force–speed gradient, and aileron wheel force. None of these parameters should be too large or too small.

The mental efforts required to fly an airplane while meeting its mission objectives are usually referred to as pilot compensation. If an airplane reacts either too slowly or too quickly to a pilot command, the pilot must compensate for this behavior by adjusting his or her own gain or by leading the airplane.

An example of pilot gain is the amount of lateral cockpit control activity generated by a pilot to keep the wings level while flying in turbulence. A high gain in this case would mean a large wheel deflection to correct for a small bank-angle deviation. A low gain would mean the opposite. *See* GAIN.

The need for a pilot to lead an airplane may be illustrated by the example of a pilot flying an airplane with a high rolling moment of inertia. When the pilot moves the lateral cockpit controls to start a turning maneuver, the airplane will be slow to react to that input. However, once this airplane begins to roll, it will be equally reluctant to stop rolling once the bank angle required for a given turning maneuver is attained. The pilot must therefore neutralize the lateral cockpit controls well before the desired bank angle is reached. In other words, the pilot must lead the airplane.

A pilot should not have to excessively lead an airplane or have to provide either too much or too little gain during any given mission segment.

Equations of motion. To predict the dynamic response characteristics of an airplane in flight and thus assess the flying qualities requires a mathematical model that describes the motions of the airplane. Since the airplane in flight has three translational and three rotational degrees of freedom (Fig. 1), it takes six equations to describe the airplane motions. In nearly all cases it is possible to describe the dynamic response behavior of an airplane by assuming that the deviations (perturbations) of the airplane motion variables from a known steady-state (constant forward speed, U), straight-line flight condition are small. This is referred to as the small-perturbation assumption. As a result of this assumption, the equations of motion can be cast in the form of six linear differential equations with constant coefficients. The six equations in turn can be split into two sets of three equations each: the longitudinal equations (drag, lift, and pitching-moment equations) and the lateral-directional equations (side-force, rolling-moment, and yawing-moment equations) respectively. *See* DIFFERENTIAL EQUATION; PERTURBATION (MATHEMATICS).

Longitudinal open-loop response. One example of the motion characteristics which result from this mathematical model concerns the longitudinal motions of the airplane. It will be assumed that the pilot moves the longitudinal cockpit control (which, in turn, moves either the elevator or the stabilizer) in a step fashion. The airplane responds to this control input with small oscillations in speed, angle of attack, and pitch attitude angle. [The speed is the true airspeed of the airplane, that is, the velocity of the airplane relative to the air. The angle of attack is the angle between the speed or velocity vector and the X axis. The pitch attitude angle is the angle between the X axis and the horizon.] *See* ELEVATOR (AIRCRAFT); STABILIZER (AIRCRAFT).

Each oscillation itself is broken down into two suboscillations: one is characterized by a low frequency, and the other by a relatively high frequency. The low-frequency motion is referred to as the phugoid mode of motion of the airplane, while the high-frequency motion is referred to as the short-period motion. In most cases, from a flying-quality viewpoint, the short-period motion is the more important. If the short-period frequency is too low, pilots call the airplane sluggish. If the short-period frequency is too high, the airplane reacts too quickly.

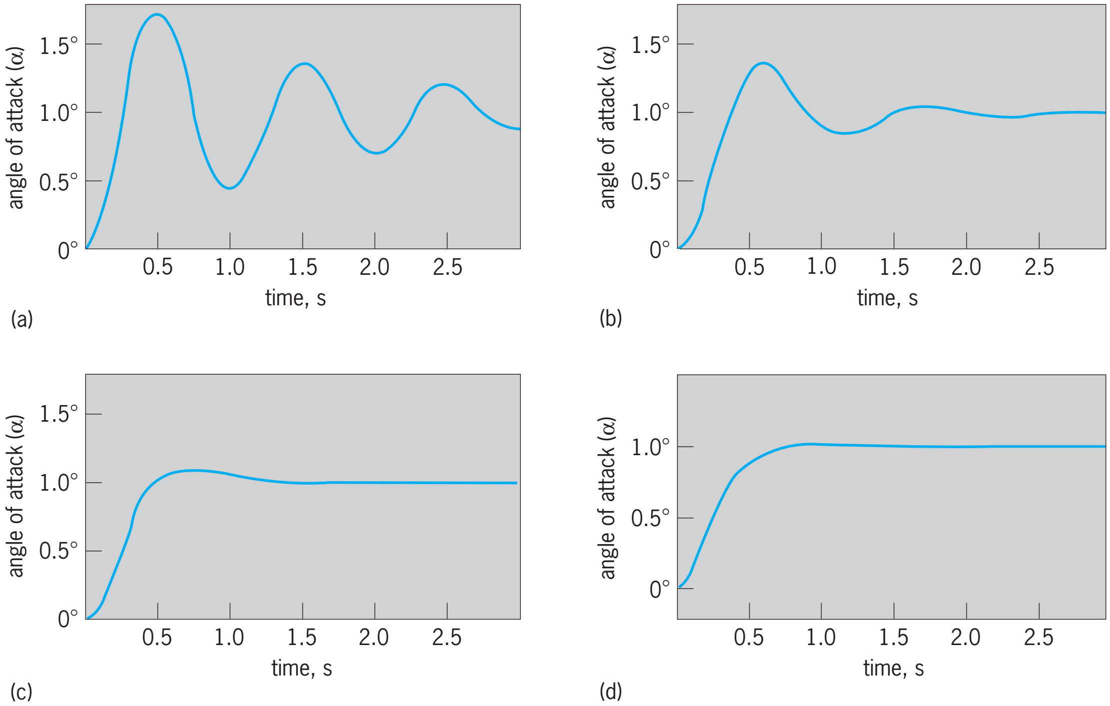


Fig. 5. Effect of short-period damping ratio. ($\zeta_{\text{short period}}$) on angle-of-attack oscillations. Typical oscillations are shown for $\zeta_{\text{short period}}$ equal to (*a*) 0.1, (*b*) 0.3, (*c*) 0.6, (*d*) 0.9.

For good control during various tracking tasks, the short-period frequency must lie between some maximum and minimum allowable values. The damping of the short period must also be acceptable (**Fig. 5**). For a damping ratio of 0.9 the airplane does not exhibit oscillatory behavior, but for a damping ratio of 0.1 there is much oscillatory activity. *See* DAMPING; MODE OF VIBRATION; VIBRATION.

Certain factors determine the short-period frequency and damping of an airplane. The short-period frequency may be estimated from factors that include the airplane lift-curve slope (the tangent of the angle between the lift and the angle of attack), the airplane center-of-gravity location, the airplane aerodynamic center location, the atmospheric density, the speed of the airplane, the wing area, and the pitching moment of inertia of the airplane. The frequency of the short period depends on the distance between the airplane aerodynamic center and the center of gravity, on the atmospheric density, and on the pitching moment of inertia. *See* MOMENT OF INERTIA.

At a forward center-of-gravity position, the airplane tends to have a higher frequency than at an aft center-of-gravity position. Also, at high altitude the frequency is significantly lower than at sea level, and airplanes tend to become sluggish. Finally, if the airplane configuration is such that large concentrated masses exist at the nose or tail (such as engines, propellers, radars, and jamming pods), the airplane would have a large pitching moment of inertia, and that also would tend to make the airplane sluggish.

The damping of the short-period oscillation is much affected by the size and the moment arm of the horizontal tail. Because the undamped natural frequency is itself proportional to the square root of atmospheric density, the damping ratio is also. Therefore, the damping ratio will deteriorate with altitude. For this reason, nearly all high-altitude airplanes have a pitch damper in their autopilot.

Closed-loop response. A typical closed-loop response scenario concerns the airplane yaw damper (**Fig. 6**). It is assumed that the airplane experiences

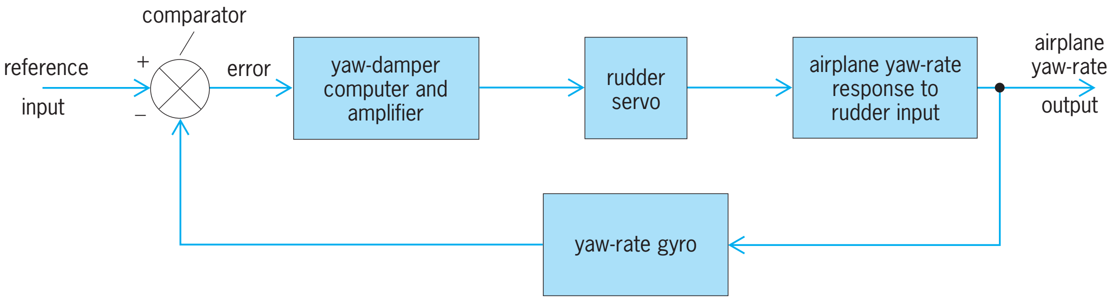


Fig. 6. Block diagram for a yaw damper.

a disturbance in the yaw rate, the angular rotation rate about the Z axis (Fig. 1). Such a disturbance could be induced by atmospheric turbulence. The yaw rate is sensed by the yaw-rate gyro and sent to a device known as the comparator. The comparator subtracts the sensed yaw-rate signal from the reference input signal of the yaw damper. In steady-state straight-line flight, this reference input signal would be zero. The output of the comparator is called the error signal, and it is sent to the yaw-damper computer. The computer decides how much rudder deflection is required and in which direction the rudder is to be moved. That signal is amplified and sent to the rudder actuator. The airplane responds to the rudder movement with a yaw rate which will tend to cancel the original yaw-rate disturbance. The yaw-damper functions to drive the error signal to zero and thereby assures that the airplane yaw rate output is equal to the reference input, which is the desired result.

This type of response scenario is called closed-loop because a loop is closed in the block diagram of the yaw damper (Fig. 6) around the airplane open-loop yaw rate response to a rudder input. The yaw damper is also called a stability-augmentation system. Its primary function is to enhance the damping of the dutch roll, which in many modern airplanes is inherently poor. *See* STABILITY AUGMENTATION.

Closed-loop systems are used to control airplanes automatically in many cases. Automatic navigation, automatic approach, and landing are typical applications. Many automatic systems can control an airplane more precisely and safely than pilots can. To ensure adequate safety in the case of system failures, many closed-loop systems are made redundant so that the failure of one or more systems is of little or no consequence. *See* AIRPLANE; AUTOMATIC LANDING SYSTEM; AUTOPILOT; CONTROL SYSTEMS; FLIGHT CONTROLS.

Jan Roskam

Bibliography. B. Etkin and L. D. Reid, *Dynamics of Flight-Stability and Control*, 3d ed., 1996; J. Roskam, *Airplane Design*, pt. 1–8, 2d ed., 1984, reprint 2000–2003; J. Roskam, *Airplane Flight Dynamics and Automatic Flight Controls*, pt. 1 and 2, 1979, reprint 2001, 2003; D. Stinton, *The Design of the Aeroplane*, 2d ed., 2001.

Flight controls

The devices and systems which govern the attitude of an aircraft and, as a result, the flight path followed by the aircraft. Flight controls are classified as primary flight controls, auxiliary flight controls, and automatic controls. In the case of many conventional airplanes, the primary flight controls utilize hinged, trailing-edge surfaces called elevators for pitch, ailerons for roll, and the rudder for yaw (**Fig. 1**). These surfaces are operated by the human pilot in the cockpit or by an automatic pilot. In the case of vertically rising aircraft, a lift control is provided. *See* AILERON; AIRCRAFT RUDDER; ELEVATOR (AIRCRAFT); ELEVON.

Controls to govern the engine power and speed, while not usually classified as flight controls, are equally important in the overall control of the aircraft. This is especially true if the engine exhaust can be directed to produce pitch or yaw motions.

Auxiliary flight controls may include trimming devices for the primary flight controls, as well as landing flaps, leading-edge flaps or slats, an adjustable stabilizer, a wing with adjustable sweep, dive brakes or speed brakes, and a steerable nose wheel.

Automatic controls include systems which supplement or replace the human pilot as a means of controlling the attitude or path of the aircraft. Such systems include automatic pilots, stability augmentation systems, automatic landing systems, and active controls. Active controls encompass automatic systems which result in performance improvement of the aircraft by allowing reductions in structural weight or aerodynamic drag, while maintaining the desired integrity of the structure and stability of flight.

Cockpit controls. The control system incorporates a set of cockpit controls which enable the pilot to operate the control surfaces. Because of the approximately fixed size and strength of the human pilot and the need to standardize the control procedures for airplanes, the primary controls are similar in most types of airplanes. The cockpit controls incorporate a control stick which operates the elevators and ailerons, and pedals which operate the rudder. Sometimes a column/wheel arrangement is used to operate the elevators and ailerons, respectively. The control stick or column/wheel is mounted on the floor between the pilot's legs. Some fighter and transport airplanes have the control stick mounted on the side of the cockpit, just beyond the end of the pilot's armrest. This space-saving arrangement, called a side-stick controller, can be selected for a variety of reasons, including improved handling qualities during maneuvering air combat and prevention of display obscuration by the control stick or column/wheel.

The longitudinal or pitching motion of the airplane is produced by fore-and-aft motion of the control stick. Forward motion of the control produces a nose-down moment of the airplane, and rearward motion a nose-up moment. Right or clockwise motion of the stick or wheel produces rolling moments to the right, and left or counterclockwise motion produces moments to the left. Moving the right rudder pedal forward produces a yawing moment to the right, and vice versa. A forward angular rotation of the rudder pedals may be used to control wheel brakes for ground control. In cases in which a lift control is provided, upward motion of the lift control lever causes increased lift.

The cockpit controls for auxiliary control devices are not as completely standardized as those for the primary controls. A nose-wheel steering control is provided on large airplanes (such as commercial

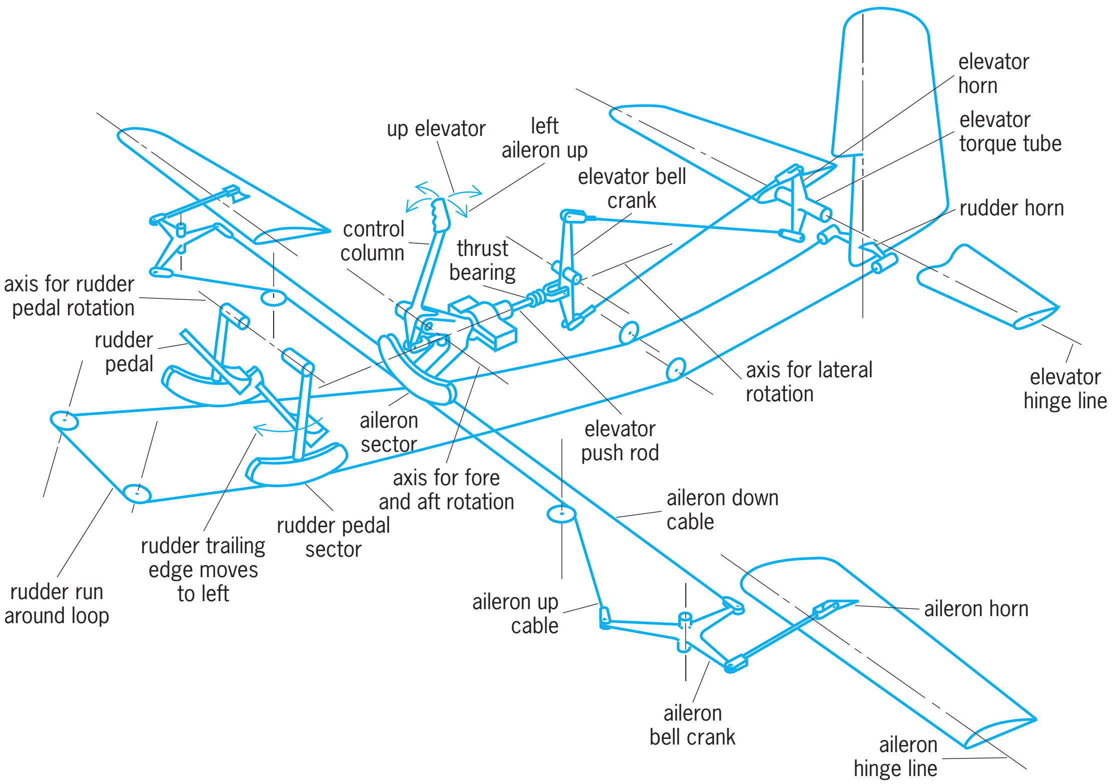


Fig. 1. Simplified diagram of a typical primary flight-control system, governing attitude and direction. (*After C. H. Chatfield, C. F. Taylor, and S. Ober, The Airplane and Its Engine, 5th ed., McGraw-Hill, 1949*)

airliners) or on airplanes that must taxi in a limited ground area (such as an aircraft carrier). The nosewheel steering may be commanded by a dedicated control wheel located on the side of the cockpit or by using the rudder pedals in combination with a switch. Trim controls, which are used to reduce the control forces to zero in a steady flight condition, may be operated by a set of control wheels or by buttons or switches which operate electric motors. In general, the directions of actuation are made consistent with those of the primary controls.

Types of control systems. Control systems with varying degrees of complexity are required, depending on the size, speed, and mission of an airplane (**Fig. 2**). In relatively small or low-speed airplanes, the cockpit controls may be connected directly to the control surfaces by cables or pushrods, so that the forces exerted by the pilot are transmitted directly to the control surfaces (Fig. 2*a*). In large or high-speed airplanes, the forces exerted by the pilot may be inadequate to move the controls. In these cases, either an aerodynamic activator called a servotab or spring tab (not illustrated) may be employed, or a hydraulic activator may be used (Fig. 2*b*). Hydraulic actuators are universally used on the controls of supersonic airplanes, because of the ineffectiveness of tabs at supersonic speeds and because of the large and sometimes irregular variations of the hinge moments on the control surfaces. In some airplanes, particularly those with swept wings and those which fly at high altitudes, the provision of adequate static stability and damping of oscillations by means of the inherent aerodynamic design of the airplane becomes difficult. In these cases, stability augmentation systems are used. These systems (Fig. 2*c*) utilize sensors such as accelerometers and gyroscopes to sense the motion of the airplane. These sensors generate electrical signals which are amplified and used to operate the hydraulic actuators of the primary control surfaces to provide the desired stability or damping. *See* STABILITY AUGMENTATION.

Auxiliary servomechanisms. Small auxiliary hydraulic or electric servomechanisms are often used to feed the inputs from the sensors to the primary control system. These devices, called secondary actuators, may be linked to the control system in such a way that no effects are felt in the cockpit, for example, by means of an extensible link or a summing linkage in the control system (a so-called series servo), or they may exert a force on the control system that causes movement of the cockpit controls (a so-called parallel servo). In some cases, the provision for both electrical and mechanical inputs is made within the package of the primary actuator. *See* SERVOMECHANISM.

Stability augmentation system failure. In the case of failure of the stability augmentation system, the human pilot should be able to override the system and maintain control of the airplane. In many airplanes, this condition is provided by limiting the travel of a series servo, or the force output of a parallel servo,

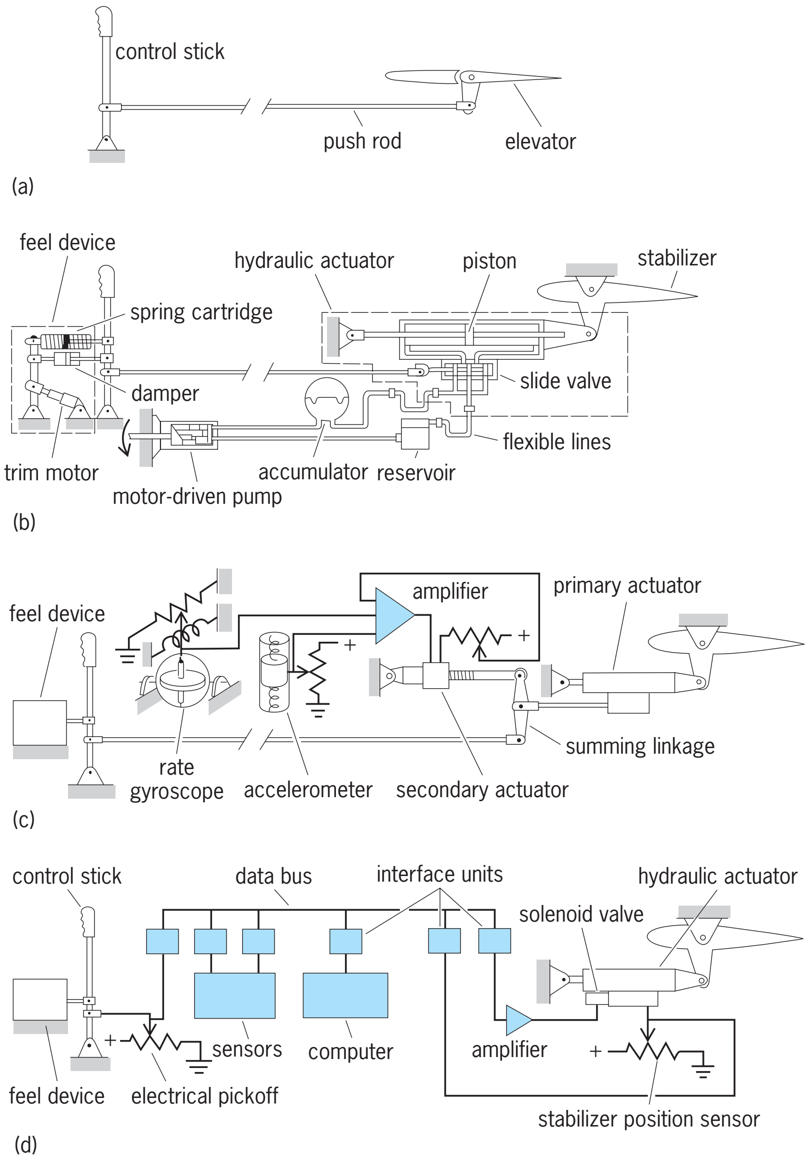


Fig. 2. Typical airplane control systems of varying degrees of complexity, illustrated for the case of longitudinal (elevator or stabilizer) control. Similar systems are used for the other control surfaces. (*a*) Mechanical control system. (*b*) Mechanical control system with control feel device and hydraulic actuator. (*c*) Control system with stability augmentation. (*d*) Digital fly-by-wire control system.

to a small fraction of that provided by the human pilot. This technique, called limiting the authority of the control, may not be acceptable if the automatic control system performs functions which require large control displacements. However, if the automatic system has large authority, its failure might result in an abrupt, large control deflection which would impose loads sufficient to cause structural failure of the airplane. The only way to avoid this problem is to make the automatic system as reliable as the airframe itself.

Fly-by-wire systems. The weight and complication of mechanical control linkages and the extensive reliance on electrical signals in automatic controls led to the development of control systems in which the control inputs from the pilot, as well as those from the stability augmentation sensors, are transmitted to the primary control actuators by electrical signals. Systems of this type are called fly-by-wire systems (Fig. 2*d*). The electrical signals are readily compatible with computers, typically digital, which can perform the functions of combining the signals from the pilot and the sensors. *See* DIGITAL COMPUTER.

Fly-by-light systems. Fly-by-wire systems can malfunction when exposed to high-intensity electromagnetic fields. The malfunctions occur when the electromagnetic fields interfere with the flight control signals being transmitted via wires between the digital computers, sensors, and actuators. The solution to this problem has been to shield the transmission media and extensively test the system before certifying it, adding cost and weight to the system. These difficulties, the intrinsic immunity of optical technology to electromagnetic interference, and the availability of optical-fiber-based transmission media from the communications industry led to the development of fly-by-light systems. These systems use optical fibers to transmit signals instead of wires; the interface units (Fig. 2*d*) are replaced with optical-electrical converters. Ultimately, the entire system, including sensors, actuators, and computers, is likely to be based on optical technology. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Digital flight control computers. There are three major reasons why digital flight control computers are used in modern airplanes. First, digital flight control computers can enhance the pilot's control of the airplane by optimizing the movement of the control surfaces in response to pilot commands, over the operating flight conditions of the airplane. Second, as a result of their ability to rapidly monitor and interpret multiple sensor inputs, digital flight control computers can often exceed the performance of an unassisted pilot in compensating for critical situations which might otherwise result in loss of airplane control. Third, digital flight control computers permit input directly from remote control and navigation devices such as digital automatic landing systems, assisting the pilot in zero-visibility conditions or freeing the pilot for other airplane management tasks.

Since digital flight control computers have a critical role in flight safety, they have stringent performance and reliability requirements. Typical requirements include a probability of failure of 10^{-9} over 10 h (for commercial airplanes); an instruction cycle time on the order of 100 Hz; and a transport lag of about 5 milliseconds. Methods of achieving the performance requirements include the use of multiple, redundant processors, fault-tolerant architectures, and parallel multitask processing techniques in both hardware and software. Control systems that make use of digital flight control computers often consist of multiple, physically separated computers with different power sources to reduce the impact of single-point failures. For fully fly-by-wire systems, triple redundancy with a fail-safe backup is considered adequate. This level of redundancy permits two failures while remaining in operation and a third simultaneous failure while maintaining the safety of flight. This design philosophy is carried through from the input sensors to the control surface actuators and their positional feedback sensors. *See* FAULT-TOLERANT SYSTEMS; MULTIPROCESSING.

Since the inputs to a digital flight control computer are usually analog at the source, they must be converted to digital signals for use by the computers. Three types of signal input are used by the flight control computers: information about the flight status of the airplane and its operating environment; information about the control surface positions; and attitude commands from the pilot or from other airplane systems, such as navigation, weapons control, terrain following/avoidance, autopilot, or automatic landing systems. *See* ANALOG-TO-DIGITAL CONVERTER.

The prime outputs of the digital flight control computer are command signals to the control surface actuators. These actuators are typically constructed so that digital signals from the computer are converted to analog signals. In addition to commanding the control surface actuators, digital flight control computers provide feedback to the pilot. This may be in terms of feel-force actuators on the cockpit controls or updated visual displays of the airplane's attitude and flight status. This feedback serves two purposes: it gives the pilot a sense of identity with the airplane, and it warns when the pilot is asking something beyond the capability of the airplane in terms of maintaining its structural integrity or dynamic stability. *See* DIGITAL-TO-ANALOG CONVERTER.

A digital flight control computer evaluates its inputs based on precomputed models of the airplane's expected behavior under various conditions of flight in order to produce command signals for the control surface actuators. Typically, wind-tunnel data are used to derive and verify the precomputed models used in the computer. Test flights and computer simulations are also used extensively to verify computer operation. *See* AIRCRAFT TESTING; WIND TUNNEL.

The primary calculation tasks of the digital flight control computer involve evaluation of analog-source input signals to produce the necessary motion of the control surfaces. Since this should result in smooth changes in attitude which are proportional to the input command signals, the computer must incorporate a considerable degree of selective signal filtering in its operation. An important function of these filters is to prevent oscillation in the control system which might otherwise result from time or value sampling differences between input, output, and feedback signals. Potential control system oscillations are typically caused by delays and mechanical resonance in the airplane's structure as it interacts with the atmosphere, as well as by the digital nature of the computers themselves. *See* ESTIMATION THEORY.

Handling qualities. Because the control system is the link through which the pilot flies the airplane, its characteristics have an important influence on the handling qualities, those characteristics which influence the ease and safety of flight. In the case of a simple mechanical control system (Fig. 2*a*), a desirable condition exists in that the control forces required of the pilot are directly related to the moments applied to the airplane, because both arise from aerodynamic effects. In such airplanes, much attention is directed to the detailed design of the control surface shapes to produce the desired hinge moment characteristics. In the case of hydraulically actuated controls (Fig. 2*b*), a control feel system must be provided to generate the control forces which are no longer available from the control surface hinge moments. Likewise, the stability augmentation system should be designed to provide response to the pilot's control inputs which have the desired magnitude and speed of response and which avoid excessive overshoot or oscillations. The techniques of servomechanism analysis are employed to design both the mechanisms themselves and the overall system, including the airplane, control system, and pilot modeling. *See* CONTROL SYSTEMS; FEEDBACK CIRCUIT.

Active controls. The availability of reliable, full-authority electronic control systems makes possible the use of closed-loop control systems to improve the performance of the airplane. The applications most frequently considered in order to take advantage of this capability are reduced static stability, maneuver load control, gust load alleviation, and structural mode damping (**Fig. 3**).

Reduced static stability. The operation of the elevators or rudder in such a way as to augment the stabilizing moment normally provided by a fixed surface (Fig. 3*a*) may allow the areas of the horizontal or vertical stabilizing surfaces to be reduced. The reduced tail area results in decreased aerodynamic drag on the airplane. In addition, in the case of the horizontal tail, a more rearward center of gravity position may be allowable, thereby reducing the down load on the tail required for trim in steady flight or in turns. The extent to which the area of stabilizing surfaces may be reduced is determined by conditions requiring full deflection of the controls, such as raising the nose for takeoff, recovery from stall, or flight with asymmetric power. Careful attention to design features which reduce these requirements allows greater benefits from active controls.

Maneuver load control. The ailerons, in addition to being deflected asymmetrically to produce roll, can be moved symmetrically upward in a turn or pull-up, to reduce the load on the wing tips (Fig. 3*b*). As a result, the bending moment on the wing is reduced, allowing a lighter structure. Alternatively, the wingspan may be increased to reduce the drag. *See* WING STRUCTURE.

Gust load alleviation. Gust load alleviation (Fig. 3*c*) may likewise be provided by moving the ailerons, or possibly the wing flaps, symmetrically to offset the load on the wing caused by gusts. The surfaces may be moved in response to the signals from an

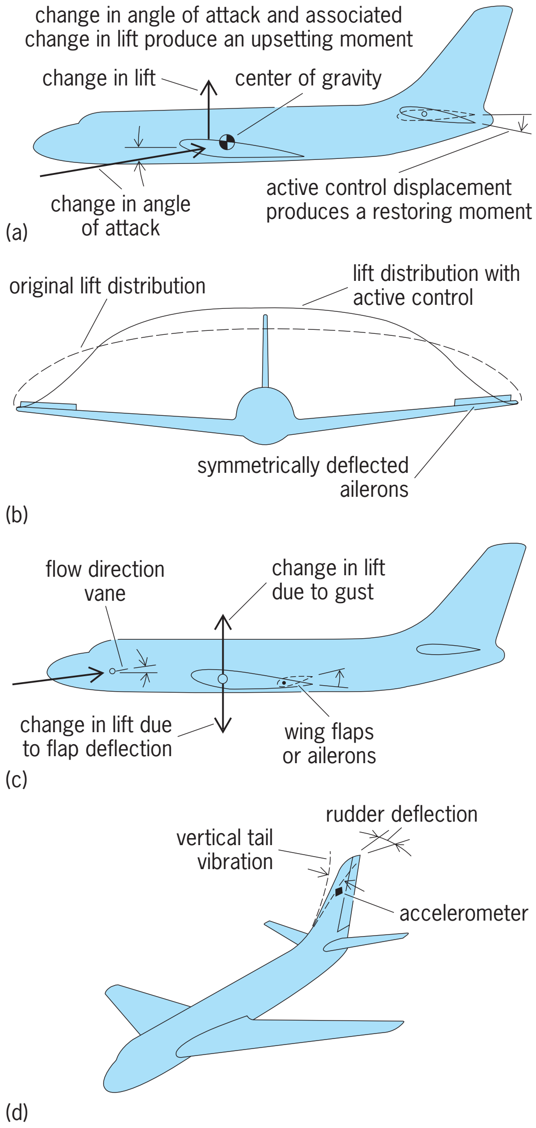


Fig. 3. Examples of active control. (*a*) Active control to allow reduced static stability. Change in angle of attack and associated change in lift produce an upsetting moment. (*b*) Active maneuver load control. (*c*) Active control for gust load alleviation. (*d*) Active control for structural mode damping.

angle-of-attack or flow-direction vane (Fig. 3*c*) or an accelerometer. The reduced gust loads may allow lighter wing structure, as well as improve passenger comfort.

Structural mode damping. Structural mode damping refers to the damping of structural vibrations caused by turbulence (Fig. 3*d*). These repeated loads are largely responsible for reducing the fatigue life of an airplane. The structural weight may therefore be reduced if these loads are decreased. In practice, the oscillations are sensed by accelerometers or rate gyroscopes at various locations in the airplane, and the control forces to damp the oscillations are applied by the primary control surfaces or by small auxiliary surfaces.

For example (Fig. 3*d*), the rudder can be moved in response to signals from an accelerometer in the vertical tail to damp lateral vibrations in the tail assembly. In more extreme applications of this technique, the flutter speed of the airplane may be increased—giving the designer greater freedom to use a lighter structure or a thinner surface, or to redistribute items of mass in a way to improve performance. *See* FLUTTER (AERONAUTICS).

Automatic flight controls. The flight of an aircraft may be controlled automatically by providing the necessary signals for navigation as inputs to the control system. In practice, automatic pilots are used to relieve the human pilot of routine flying for long periods, and automatic control systems are used to make precision landings and takeoffs under conditions of reduced visibility. The reference signals for such navigation functions come from radio signals generated by ground stations or by satellites, and from inertial navigation systems. Inertial navigation systems determine the position of the airplane over the surface of the Earth by use of precise on-board gyroscopes and accelerometers. *See* AUTOPILOT.

The requirements for reliability of the navigation systems for cruise flight are less critical than those for the primary control system, because of the time available to recover from a failure. However, the requirements for reliability of automatic landing systems are similar to those for the primary control system. *See* AUTOMATIC LANDING SYSTEM; ELECTRONIC NAVIGATION SYSTEMS; INERTIAL GUIDANCE SYSTEM; SATELLITE NAVIGATION SYSTEMS.

Flight management. The use of a flight-control system to provide the desired responses to control inputs, to limit structural loads, and to fly on precise paths defined by inertial or external navigation references relieves the human pilot of many tasks requiring skill and dexterity. The pilot is thereby given the opportunity to concentrate on management of the systems which control the airplane. However, even this task can overwhelm the pilot. Therefore, designers are constantly searching for ways to automate airplane control functions. Applications of artificial intelligence concepts, including the use of symbolic manipulation programming languages, are resulting in the development of embedded expert systems to perform a variety of automatic control functions, including automatic selection of navigation aids and other sensors, mode control selection based on flight condition, and automatic fault detection and control system reconfiguration. Although more and more control functions are being automated, designers are careful to provide advisory displays which keep the pilot aware of the status of various sytems and to provide mechanisms which permit the pilot to override an automatic control function if necessary. *See* AIRCRAFT INSTRUMENTATION; ARTIFICIAL INTELLIGENCE; EMBEDDED SYSTEMS.

William H. Phillips; P. Douglas Arbuckle

Bibliography. J. D. Anderson, *Fundamentals of Aerodynamics*, 3d ed., 2001; J. Blakelock, *Automatic Control of Aircraft and Missiles*, 2d ed., 1991; J. Curran, *Trends in Advanced Avionics*, 1992; B. Etkin and L. D. Reid, *Dynamics of Flight Stability and Control*, 3d ed., 1996; G. F. Franklin,

J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*, 5th ed., 2006; J. Roskam, *Airplane Flight Dynamics and Automatic Flight Controls*, pts. 1 and 2, 1979, reprint 2001, 2003; C. Spitzer, *Digital Avionics Systems: Principles and Practice*, 2d ed., 1993, reprint 2000.

Floodplain

The relatively broad and smooth valley floor that is constructed by an active river and periodically covered with water during intervals of flooding. Water chemistry and suspended sediment patterns indicate that floodwaters can come from combined main channel overflow, local ground-water rise, tributary inputs, and precipitation directly onto a saturated floodplain surface. Engineers consider the floodplain to be any part of the valley floor subject to occasional floods that threaten life and property, and commonly designate floodplain zones based on the recurrence interval with which they are flooded (for example, 100-year floodplain). Various channel modifications, from levees and channelization to flow regulation, can be used to restrict the natural flood process. These modifications have altered most of the world's large river-floodplain systems. For example, an estimated 96% of the floodplain ecosystem historically present along the lower Mississippi River has been lost because of levees and channelization. *See* RIVER; RIVER ENGINEERING.

Geomorphologists consider the floodplain to be a surface that develops by the active erosional and depositional processes of a river. Floodplains are underlain by a variety of sediments, reflecting the fluvial history of the valley, and floodplain stratigraphy is commonly used to infer paleoenvironmental changes. The chronology of these changes is determined using radiocarbon, thermoluminescence, and other dating techniques. Large rivers can have extensive floodplain deposits; the Mississippi River has up to 300 ft (100 m) of fluvial sediment beneath its floodplain, and the floodplain of the Amazon River reaches 30 mi (50 km) across. *See* DATING METHODS; FLUVIAL SEDIMENTS; GEOMORPHOLOGY.

Deposits and topography. Most floodplains consist of the following types of deposits: colluvium—slope wash and mass-wasting products from the valley sides, as is common along steep channels with narrow floodplains; channel lag—coarse sediment marking the bottoms of former channels; lateral accretion deposits—sand and gravel deposited as a meandering river migrates laterally; vertical accretion deposits—clay and silt deposited by overbank flooding of the river; crevasse-splay deposits—relatively coarse sediment carried through breaks in the natural river levees and deposited in areas that usually receive overbank deposition; and channel-fill deposits—sediment filling abandoned river channels. Channel fills can be coarse for sandy rivers. Muddy rivers such as the Mississippi, however, fill their abandoned channels with cohesive clay. The clay "plugs" in these old channels then resist further migration of the river. Changes in the river through time create a complex floodplain topography (**Fig. 1**). *See* STREAM TRANSPORT AND DEPOSITION.

Fig. 1. **Aerial view of the floodplain of the Amazon River showing complexity of floodplain topography. 20 km = 12.4 mi. (*After L. A. K. Mertes et al., Channel-floodplain geomorphology along the Solimões-Amazon River, Brazil, Geol. Soci. Amer. Bull., 108:1089–1107, 1996*)**

On a section across a typical broad floodplain, such as that of the Mississippi, the highest elevations occur at the crests of natural levees flanking the river channels. The levee back slopes flatten away from the channels toward backswamp basins. These basins receive fine overbank deposits during occasional floods. Their high water tables promote irregular lakes and swamp or marsh vegetation. Levee crests can rise 33 ft (10 m) or more above adjacent backswamps along certain portions of the Mississippi. Crevasse splays through the levees can locally extend sandy deposits into backswamp areas.

The floodplain of the world's largest river, the Amazon, is annually inundated for 4 to 5 months. Water level rises 30 ft (9 m) vertically along some parts of the river, and spreads up to 30 mi (50 km) across the floodplain. Vertical change in water level as a function of discharge is related to floodplain width, which varies substantially along the length of most rivers. At the downstream end of a river that enters a body of standing water, the floodplain grades into a delta. *See* DELTA.

Floodplain functions and dynamics. Floodplains influence physical processes along rivers because the floodplain surface often has greater hydraulic roughness, associated with vegetation or structures,

Fig. 2. Flooded forest along the Amazon River. Water depths can reach several meters in areas such as this during the peak of the annual flood.

and shallower flows than the main channels. Thus floodwater moves across floodplains more slowly, dissipating flow energy and attenuating the flood hydrograph. Floodplains also strongly influence river ecological processes by providing spawning and nursery habitat for many species during floods, and nutrients to the main channel as floodwaters recede. Habitat complexity and abundance increases substantially during floods as wetlands, backwaters, flooded forests (**Fig. 2**), and floodplain lakes all become accessible to organisms migrating out from the main channel. These ecological processes are described by the flood-pulse model, which highlights the ecological necessity of regular seasonal floods along many alluvial rivers. River restoration efforts around the world now attempt to restore channel–floodplain connections through levee setbacks, changes in flow regulation, and removal of channelization measures.

Sediment can remain on a floodplain for hundreds or thousands of years, or the sediment can have a very short residence time. Residence time depends on the recurrence interval of high-energy floods capable of depositing or eroding floodplain sediments (**Fig. 3**). Laterally stable rivers can alternate between gradual floodplain construction through overbank deposition over hundreds of years, and episodic erosion that strips the floodplain to a basal lag deposit from which it slowly reforms. Laterally mobile rivers are more likely to have regular erosion and redeposition of new floodplain sediments as the river migrates. These different patterns of floodplain storage can be particularly important where contaminants such as heavy metals or organochlorine compounds are attached to silt and clay particles. In this case, floodplain dynamics determine how these contaminants are stored and mobilized along a river network. *See* SEDIMENTOLOGY; STREAM TRANSPORT AND DEPOSITION.

River terraces. Many rivers display abandoned floodplains that formed when the river flowed at higher levels than at present. These abandoned floodplains are known as terraces. Terraces reflect the numerous complex controls on river gradients and patterns, including climatic changes, uplift or subsidence of the drainage basin, sea-level changes, and land-use changes. Distinguishing between these various controls for a given field setting is often very

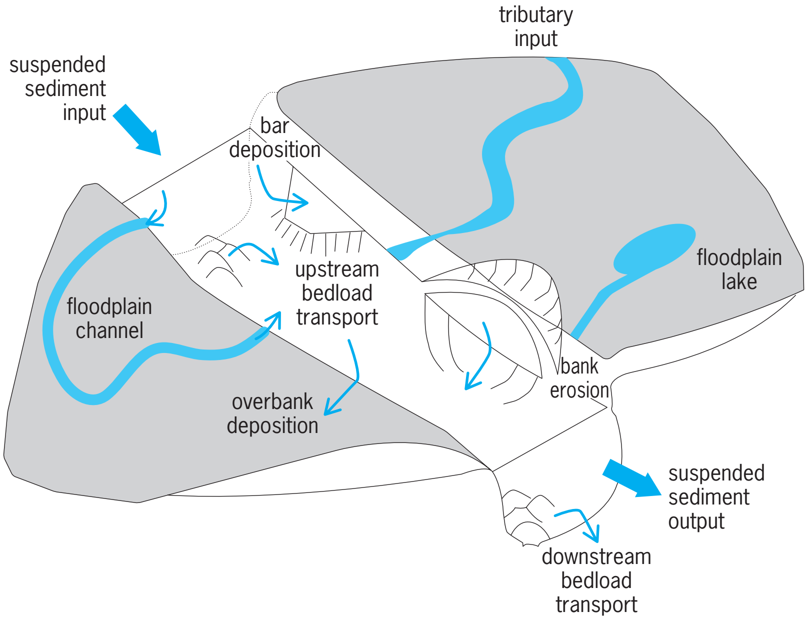


Fig. 3. Processes governing sediment storage along a channel–floodplain reach. Sediment enters a reach of channel from upstream, tributaries within the reach, and bank erosion. Sediment leaves the channel reach by channel transport, deposition on bars, diffuse overbank flow, and through floodplain channels. (*After T. Dunne et al., Exchanges of sediment between the flood plain and channel of the Amazon River in Brazil, Geol. Soc. Amer. Bull., 110:450–467, 1998*)

difficult, but has the potential to provide important insights into environmental changes and river response through time. *See* FLUVIAL EROSION LANDFORMS. Ellen E. Wohl; Victor R. Baker

Bibliography. P. B. Bayley, The flood pulse advantage and the restoration of river-floodplain systems, *Regul. Rivers Res. Manage.*, 6:75–86, 1991; P. B. Bayley, Understanding large river-floodplain ecosystems, *BioScience*, 45:153–158, 1995; J. S. Bridge, *Rivers and Floodplains: Forms, Processes, and Sedimentary Record*, Blackwell, Oxford, 2003; T. Dunne et al., Exchanges of sediment between the flood plain and channel of the Amazon River in Brazil, *Geol. Soc. Amer. Bull.*, 110:450–467, 1998; C. W. Martin, Heavy metal storage in near channel sediments of the Lahn River, Germany, *Geomorphology*, 61:275–285, 2004; L. A. K. Mertes, Inundation hydrology, pp. 145–166, in E. Wohl (ed.), *Inland Flood Hazards*, Cambridge University Press, 2000; R. E. Sparks, Need for ecosystem management of large rivers and their floodplains, *BioScience*, 45:168–182, 1995.

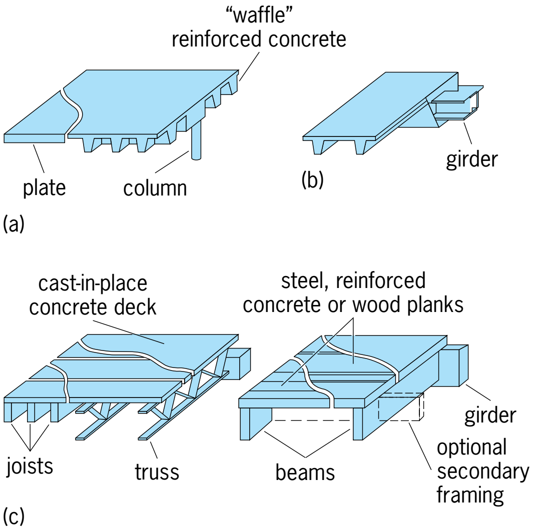
Floor construction

A floor of a building generally provides a wearing surface on top of a flat support structure. Its form and materials are chosen for architectural, structural, and cost reasons.

A ground-supported floor may be of almost any firm material, ranging from compacted soil to reinforced concrete. It is supported directly by the subsoil below.

An elevated floor spans between, and is supported by, beams, columns, or bearing walls. It is designed to be strong and stiff enough to support its design loading without excessive deflection; to provide for an appropriate degree of fire resistance; and to supply diaphragm strength to maintain the shape of the building as a whole, if necessary. A ceiling may be hung from the underside of the floor assembly as a finish surface for the room below. The optimum floor design meets these criteria while being as thin as possible and economical to construct. *See* BEAM; COLUMN; LOADS, TRANSVERSE.

Wood. Wooden floors are generally used in light residential construction. Such flooring generally consists of a finish floor installed on a subfloor of tongue-and-groove planking or plywood, spanning between wooden beams that are commonly called joists (see **illus.**). The flooring is usually about 1 in. (2.5 cm) thick; the joists are, for example, 2 in. by 8 in. (5 cm by 20 cm), spaced at 16 in. (40 cm) or 24 in. (60 cm) on center. Such distances are widely recognized in modular-type construction. The joists usually span about 10–20 ft (3–6 m) between their supports. The dimensions are determined by architectural requirements, code-stipulated design loads, and allowable design stresses for the floor components. The wearing surface of such wooden floors is typically of vinyl tiles or hardwood flooring.



Some types of framing and decking floor construction. (*a*) Cast-in place. (*b*) Precast. (*c*) Site-assembled. (*After Sweets Catalog File, General Building and Renovation, 1993*)

Reinforced concrete. Slabs fabricated of reinforced concrete are a common type of floor for heavier loading. The concrete is cast on forms, and reinforced with properly placed and shaped steel bars (rebars), so as to span between steel or reinforced concrete beams or between bearing walls. Flat slab floors are supported primarily by columns alone. Sometimes reinforced concrete planks precast in a factory are used for the slab structure, covered by a thin finish cover of a strong concrete mix that is troweled smooth. *See* REINFORCED CONCRETE.

Composite. This type of floor is commonly used in modern office building construction. Concrete is cast on, and made structurally integral with, corrugated metal deck, which spans between steel joists of either solid-beam or open-web types, generally spaced between about 16–48 in. (40–120 cm) on center. The corrugated metal deck may include closed tubular regions to provide space for electrical conduit. Sometimes such composite deck is also structurally connected to the supporting steel beams so as to make them act as composite beams and thus be stronger, stiffer, and more economical for their function. *See* COMPOSITE BEAM.

Prestressed concrete. This is used for long span slabs. Highly prestressed high-tension steel wires within the high-strength concrete slab produce a thin, stiff, and strong floor deck. *See* PRESTRESSED CONCRETE.

Lift slab. This type of construction has evolved for economy and efficiency. A concrete slab is first formed at ground level, reinforced and cured to adequate strength, and then carefully jacked up into its final position on supporting columns. *See* STRUCTURAL DESIGN. Milton Alpern

Bibliography. American Society of Civil Engineers, *Minimum Design Loads for Buildings and Other Structures*, ASCE 7–88, 2000; D. K. Ballast, *Architect's Handbook of Construction Detailing*, 1990;

H. J. Cowan, *Design of Reinforced Concrete Structures*, 2d ed., 1988; K. F. Faherty and T. G. Williamson (eds.), *Wood Engineering and Construction Handbook*, 2d ed., 1995.

Floriculture

The segment of horticulture concerned with commercial production, marketing and sale of bedding plants, cut flowers, potted flowering plants, foliage plants, flower arrangements, and noncommercial home gardening.

Florist crop production. Commercial crops are grown either in the field or under protected cultivation, such as in glass or plastic structures. Field production is confined to warm climates or to summer months in colder areas. Typical field crops are gladiolus, peonies, stock, gypsophila, asters, and chrysanthemums. Greenhouse production is not as confined by climate or season, but most greenhouses are located in areas that have advantages such as high light intensity, cool night temperatures, or ready access to market. Jet air transportation resulted in major changes in international crop production. For example, fast transport of inexpensive cut flowers from Bogotá, Colombia, to Europe and the United States greatly affected production on all three continents. Other newer crop production centers, such as Israel, could cause additional changes in the world production and distribution of cut flowers. The cost of shipping potted plants has generally confined large-scale potted plant production to population centers.

Costs. Greenhouse production and profits are calculated on a square-foot basis rather than on an acre basis as with most agricultural crops. Costs are high because of labor, expensive greenhouse structures and equipment, fuel required for heating, and handling and shipping of the perishable products. Commercial flower growers are quick to utilize research results to increase efficiency, quality, production and profit.

Soilless media. A striking change that has occurred in commercial greenhouse production is the almost universal acceptance of soilless mixes as growing media, particularly for potted plants. In some areas the most popular medium is peat moss and vermiculite, or peat and perlite, while in other areas such as the southeastern United States it is pine bark humus and sand. Many of the media are formulated by processors who sell the potting mixture to growers. Advantages of the prepared mixes can be uniformity of product from one shipment to the next, freedom from pests, and lightness. The media are also well drained with adequate pore space. Soilless media gained in popularity when field soils became less available or of questionable quality at greenhouse ranges located in or near metropolitan areas.

Controls in the greenhouse. The greenhouse environment can be controlled to such an extent that flowers can be produced on very precise schedules year-round. The phenomenon of photoperiodism is used by growers to produce chrysanthemum flowers when the natural day length is too long for flower bud initiation and development. Growers simulate short-day conditions by pulling black cloth over the plants each evening. The short-day plants can be kept in a vegetative state by interrupting the dark period with lights, when the natural day length would be conducive for flowering. Flower bud dormancy of azaleas and hydrangeas can be overcome by exposing plants to cool temperatures. Proper nutrient levels are maintained with sophisticated fertilization programs. Pests are controlled with systemic fungicides and insecticides, and biological control is often used. *See* PHOTOPERIODISM.

Growth and flowering also are regulated with chemicals. Some chemicals enhance rooting of cuttings, while other stimulate lateral shoot development, inhibit growth, stimulate flower bud initiation or development, break dormancy, or affect flower color. Perhaps the most commonly used chemicals are the growth retardants which prevent potted plants such as poinsettias and chrysanthemums from becoming too tall.

Breeding. Pronounced improvements in cultivars have been realized because of excellent breeding programs conducted by commercial propagators and by some horticulture departments. Modern cultivars have traits such as more attractive flower colors and forms, longer-lasting flowers, better growth habit, increased resistance to insects and disease organisms, or ability to grow and flower at cooler night temperatures. *See* BREEDING (PLANT).

Field production. Most environmental factors cannot be as well controlled in field production as in the commercial greenhouse, but growers follow some procedures to regulate growth and flowering. Fields can be lighted at night to keep chrysanthemums vegetative until growers wish to provide short days for flowering. Black cloth or plastic can be pulled over plants at night in the summer to promote flowering, though the expense of this practice has reduced its popularity and feasibility. Plants often are covered with Saran cloth to reduce light intensity to prevent injurious leaf and flower temperatures. Irrigation systems are installed to ensure adequate water, and soluble fertilizers can be injected through the irrigation system. Fields are fumigated prior to planting to control pests, and crops can be grown at closer spacings since cultivation is not required to control weeds. Grading and packing operations have been improved to get the flowers to market in good condition.

Wholesale marketing. The wholesaler serves as a valuable link between the grower and florist. Cut flowers and florists' greens primarily are marketed by the wholesaler, though some potted plants pass through this channel. Wholesalers usually are located in metropolitan areas, near retail florist shops, but some wholesalers use trucks to service shops quite far from the wholesale house. The wholesale florist also has a large inventory of supplies, such as ribbon, foil, containers, and other items essential in floral design.

Retail floriculture. The retail florist often is the final link in the movement of flowers and potted plants from the field to greenhouse to the customer. Florists are especially busy at holidays, but many floral arrangements and corsages are purchased for special occasions throughout the year. The florist offers various customer services, such as delivery, artistic design, and credit. An organized network of florist shops makes its possible for flowers to be ordered from one shop and delivered by another in a distant city.

Cut flowers and potted plants also are sold by mass market outlets such as supermarkets, department stores, and discount centers. Flowers and potted plants are purchased in large quantities by these outlets. Most sales are on a cash-and-carry basis, eliminating costs of delivery and billing. A large percentage of bedding plants are sold in supermarkets.

The purchase of flowers on a weekly basis is not yet as common in the United States as in Europe, but the increased interest in the environment has had pronounced effects on the floriculture industry.

Home gardening. Expanded interest in the environment has increased the number of individuals who now enjoy gardening as an avocation. Bedding plants, such as petunias, marigolds, geraniums, and impatiens, can produce pleasing effects throughout the summer at little cost. *See* PLANT GROWTH.

Roy A. Larson

Bibliography. M. N. King (ed.), *Introduction to Greenhouse Production*, 1992; R. Larson (ed.), *Introduction to Floriculture*, 2d ed., 1992; G. MacDaniel, *Ornamental Horticulture*, 1991.

Flotation

A process used to separate particulate solids, which have been suspended in a fluid, by selectively attaching the particles to be removed to a light fluid and allowing this mineralized fluid aggregation to rise to where it can be removed. The principal use of the process is to separate valuable minerals from waste rock, or gangue, in which case the ground ore is suspended in water and, after chemical treatment, subjected to bubbles of air. The minerals which are to be floated attach to the air bubbles, rise through the suspension, and are removed with the froth which forms on top of the pulp. Although most materials subjected to flotation are minerals, applications to chemical and biological materials have been reported.

Fields of application. Froth flotation was first used to recover sulfide minerals which were too fine to be recovered by gravity concentrators such as jigs, tables, and sluices. These early operations removed all of the sulfide minerals into one concentrate. Although good recoveries were achieved by this bulk flotation, the concentrate grades were poor.

As the understanding of the chemistry of flotation improved, better reagents became available and differential flotation became a reality. It became possible, in many cases, to float only the valuable sulfide minerals and leave the unwanted pyrite (fool's gold) with the gangue. The advent of differential flotation revolutionized the mineral-processing industry. Most plants discarded their gravity-concentrating equipment and turned to flotation exclusively.

Froth flotation is also used to separate the silicate minerals. It is quite common to produce a concentrate of albite ($NaAlSi_3O_8$), muscovite $[KAl_2(AlSi_3O_{10})(HO_2)_1]$, quartz ($SiO_2$), spodumene ($LiAlSi_2O_6$), and zircon ($ZrSiO_4$) by flotation.

Flotation is also used to concentrate oxide minerals such as hematite (Fe_2O_3) and pyrolusite (MnO_2), and native elements such as sulfur, silver, gold, copper, and carbon (both graphite and diamond), as well as many salts. Slightly soluble salts such as fluorite (NaF), barite ($BaSO_4$), and calcite ($CaCO_3$) can be floated with conventional flotation techniques, but soluble salts such as halite ($NaCl$) and sylvite (KCl) are floated not from water but from saturated brines.

Diamonds and phosphates are floated by immersing the ore in water on which there is a thin layer of oil. The gangue sinks, and the diamonds (or phosphates) float on the water with the oil. This type of flotation is known as skin flotation.

Flotation has been widely used in the coal-cleaning industry; the coal is floated, and the pyrite and shales, which form ash, are depressed into the pulp.

Principles. The reagents used in flotation can be classified as collectors, frothers, and modifiers.

Collectors. Collectors are organic molecules which have two distinct parts: polar and nonpolar. The polar end of the collector molecule attaches to the mineral surface; the nonpolar portion of the molecule is exposed to the water. The thin coating of collector transforms the minerals from hydrophylic to hydrophobic. When the coated mineral particle comes into contact with an air bubble, the particle adheres to the bubble and rises through the pulp, as shown in **Fig. 1**. Collectors are classified as either cationic or anionic. Examples of anionic collectors are xanthates, sulfonates, thiophosphates, and organic acids. The principal cationic collectors are amines.

The mechanism of adsorption of collectors onto mineral surfaces and the resulting bubble attachment has been debated. Adsorption is controlled by a combination of driving forces, any one of which may dominate in any given case. Among the recognized driving forces are: chemical reaction between the mineral and the collector; attraction of opposite charges between the mineral surface and the polar portion of the collector; and association between the nonpolar collector tails. *See* ADSORPTION.

Frothers. Frothers are added to the mineral pulps to help form a stable froth which can be removed. Frothers perform this function by becoming more concentrated at air-water interfaces than in the body

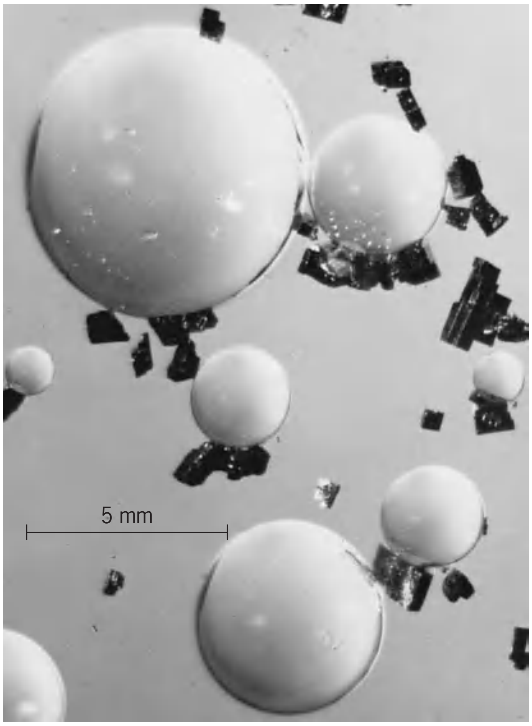


Fig. 1. Photomicrograph of particle-bubble attachment in flotation. Particles of xanthate-conditioned 65/100-mesh galena are clinging to air bubbles and are being rafted to the surface. View is horizontal through the side of a glass flotation cell. (*Courtesy of W. R. Speden*)

of the water. This changes the surface free energy of the air-water interface. When a bubble strikes the pulp-air interface, it does not break, because the kinetic energy associated with its movement can be adsorbed by diffusion of more frother to the interface. Because the surface must absorb the energy, it is important that the concentration of frother be such that a small change in frother concentration produces a large change in surface energy. For this reason, an excess of frother can "kill" froth.

Frothers, somewhat like collectors, tend to be heteropolar molecules, with a polar and nonpolar portion. Some collectors, such as oleic acid and the amines, also produce a froth. It is important that the frother not act as a collector and float an unwanted mineral.

The chemicals used as frothers are primary alcohols, secondary alcohols, and ethers.

Modifiers. Modifiers are added to flotation systems to alter the chemical behavior of the collector or the mineral, or both. If cyanide (CN^-) is added to a pulp which contains pyrite (FeS_2), an iron-cyanide complex is formed on the surfaces of the pyrite which prevents the attachment of the collector and, therefore, prevents the flotation of the pyrite.

Another example is the addition of copper ion (Cu^{2+}) to activate sphalerite (ZnS). Xanthates make excellent collectors for copper sulfide minerals, but poor collectors for zinc minerals. When copper ions are added to the pulp, there is an exchange of zinc for copper at the surface of the sphalerite which causes the surface of the zinc mineral to respond, chemically, as if it were a copper mineral, and hence it can be easily floated with the xanthate collectors.

The interaction of the air bubble and particle in the pulp, and the subsequent attachment, appears to be a dual process. The pulp is supersaturated with gas (air), and a bubble nucleates on the collector-coated mineral. The collision of this particle with a free-moving air bubble results in an easy attachment of the two.

Plant practice. Commercial-scale flotation is carried out in a continuous operation with equipment similar to that shown in **Fig. 2**. These cells can be as small as $^1/_3$ ft^3 to as large as 500 ft^3 (14 m^3).

Among the operational variables are the pulp density (ratio of solids to liquid), the amount of reagent added, the time that the pulp spends in the cells, the height of the froth, the speed of the agitator, the temperature of the pulp, and the size of the solids in the pulp. There are also several variables associated with each ore which cannot be controlled. In light of these variables, it is safe to say that no two flotation plants have the same flow sheet or reagent schedule.

There are many approaches which can be taken to separate a mixture of valuable minerals. One is to float all of the valuable minerals together and then separate them. This approach is used to recover chalcopyrite ($CuFeS_2$) and molybdenite (MoS_2). Both minerals are floated; xanthates are used to collect the chalcopyrite, and fuel oil is used to activate the molybdenite. It should also be noted that cyanide and lime [$Ca(OH)_2$] are added to the pulp to prevent the

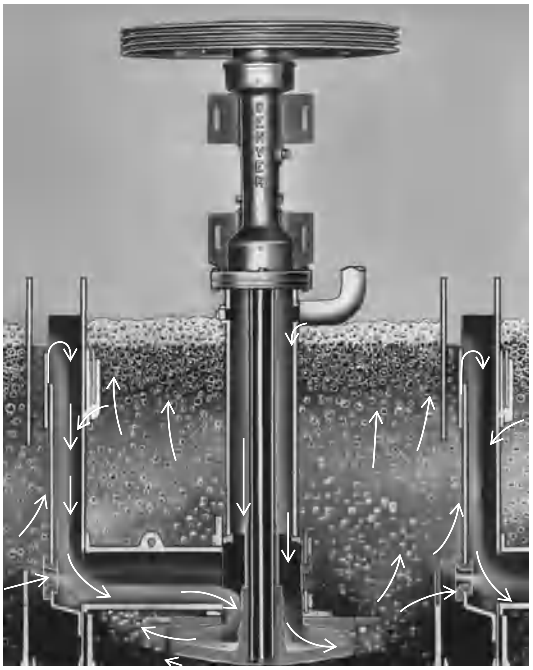


Fig. 2. Cutaway view of section of a flotation cell. (*Denver Equipment Co.*)

pyrite from floating. The chalcopyrite-molybdenite concentrates are then treated with a polysulfide, known as Noke's reagent, which prevents the chalcopyrite from floating. The next stage of flotation produces a molybdenite concentrate in the froth and a chalcopyrite concentrate in the tails.

This separation can also be approached by floating each mineral separately. The molybdenite is floated, the xanthate can be added to the pulp, and the chalcopyrite recovered. This approach works well, but a plant almost twice as large as that used for the bulk float approach described above is needed.

Separating some ores requires both approaches. An example is the separation of a mixture of chalcopyrite ($CuFeS_2$), galena (PbS), sphalerite (ZnS), and pyrite (FeS_2) in a calcite gangue. It is easy to float the chalcopyrite and the galena together and then separate them. After the galena and chalcopyrite are removed, the sphalerite can be floated.

Trends in flotation have been toward using larger cells with automatic equipment to control such variables as pH, pulp density, and froth height. The amount of material treated by flotation is enormous. Plants which can handle in excess of 100,000 tons (90,000 metric tons) per day are common. *See* INTERFACE OF PHASES; ORE DRESSING; SURFACTANT; UNIT OPERATIONS.

Ronald L. Atwood

Bibliography. R. D. Crozier, *Flotation*, 1992; M. C. Furstenau and J. D. Miller, *Chemistry of Flotation*, 1985; M. W. Ranney (ed.), *Flotation Agents and Processes: Technology and Applications*, 1981.

Flow-induced vibration

Structural vibrations caused by a flowing fluid. Whenever a structure is exposed to a flowing fluid, the fluid forces may cause the structure to vibrate. In many cases, such as a gentle wind rustling the leaves of a tree, the structural response does not produce any damage. However, in other cases, the structural motion can couple with the fluid excitation in such a way that the energy from the flowing fluid is transferred to the structure and damaging vibrations can occur. One of most spectacular examples of this is the 1940 collapse of the Tacoma Narrows Suspension Bridge. In this case, the 2800-ft (850-m) central span oscillated wildly for several hours before finally failing at a relatively modest wind speed just over 40 mi/h (65 km/h). In fact, a similar catastrophe had occurred nearly 100 years earlier with the Wheeling Suspension Bridge in Ohio. This bridge had a span of 1000 ft (305 m), the longest in the world at the time (1854). Any lessons learned from this experience had apparently long been forgotten when the Tacoma Narrows Bridge was designed. Fortunately, not all excessive flow-induced vibrations lead to failure if they are recognized and appropriate action is taken. The Golden Gate Suspension Bridge was completed in 1937 and has a central span of 4200 ft (1280 m). During a wind storm in 1951, the bridge deck suffered peak-to-peak vibration amplitudes of 12 ft (3.5 m) and peak-to-peak rotational oscillations of 22°. Subsequent torsional stiffening of the deck stabilized the bridge response to acceptable levels. *See* BRIDGE; FLUID MECHANICS; MECHANICAL VIBRATION; VIBRATION.

Impact of changing design approaches. Traditionally, structures were designed with large factors of safety, which were required to compensate for the effects of limited knowledge of loading and strength as well as for the relatively simple analysis tools available at the time. The resulting structures were massive and stiff, and they did not respond significantly to fluid loading. Even when light weight was a design criterion, the structures were stiff, such as is seen in the early biplane and triplane designs. However, the introduction of the monoplane for increased performance also substantially reduced wing stiffness. Thus, the wings of the Fokker D8 were found to oscillate abruptly at a certain wind speed and to break off after a few oscillations. The flexibility of the wing permitted bending and twisting vibrations that coupled with the wind to produce a hitherto unknown flow-induced vibration phenomenon. This phenomenon has become known as bending-torsion flutter, and it is a critical design criterion for modern fixed-wing aircraft. *See* AEROELASTICITY; FLUTTER (AERONAUTICS); STRUCTURAL ANALYSIS.

The technology explosion since the mid-twentieth century has revolutionized our understanding of materials and structural loading as well as our analysis tools. This has permitted a much improved approach to design and the establishment of rational factors of safety. These advances, on top of the demand for higher efficiency and performance, have resulted in lighter, more flexible structures exposed to higher-velocity flows. All of these factors have increased the likelihood of flow-induced vibrations. The problems are ubiquitous, from the fluttering of artificial heart valves and the galloping of bridge decks to the oscillations of marine structures in waves and tidal currents and damaging vibrations of tubes in nuclear steam generators. Unfortunately, the fields of application are so wide ranging that the literature is quite scattered, tending to be focused in specialized areas of application such as aeronautics, wind engineering, ocean engineering, hydraulics engineering, nuclear engineering, and biomechanics. Because of the particular difficulties in analyzing problems of dynamic instability and unsteady fluid flows with moving boundaries, substantial research has been carried out in the fields of theoretical mechanics and computational fluid dynamics. The latter is especially challenging and a field of much current research activity. *See* COMPUTATIONAL FLUID DYNAMICS.

Basic mechanisms. Despite the great diversity of the physical applications of flow-induced vibrations, they can generally be classified into three basic mechanisms: forced vibrations, self-excited vibrations, and self-controlled or resonant vibrations. These distinguish the phenomena both physically and mathematically and can be illustrated generically by plotting structural response amplitude against flow velocity

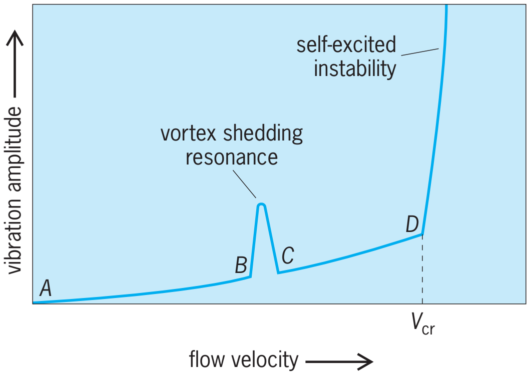


Fig. 1. Generic vibration amplitude response as a function of flow velocity.

(**Fig. 1**). The details of the response characteristics are discussed separately with each excitation mechanism below.

Forced vibration. Forced vibration phenomena are characterized by fluid excitation forces which do not interact strongly with the structural motions; that is, the fluid forces are not affected by the structural response and contain no significant periodicities that could excite structural resonance. This is the simplest form of flow-induced vibration.

Mathematically, such problems involve excitation forces that are functions of time alone, and the solution represents a steady-state equilibrium between the energy input to the structure from the flow and the energy dissipated by structural damping. As the flow velocity increases, the energy in the flow increases and so too will the structural response (Fig. 1, regions A–B and C–D).

Typically, modeling the structural dynamics is fairly straightforward and uses empirical damping factors. Unfortunately, modeling the fluid mechanics in many cases is much more challenging and measured data are usually required. However, since the structural motion does not affect the fluid forces, the experimental determination of the forces is relatively simple and can be obtained from rigid scaled wind-tunnel models or from measurements on full-scale structures. *See* WIND TUNNEL.

The most common example of this type of flow-excitation mechanism is the turbulent buffeting of structures by the wind. The wind near the Earth's surface behaves like a turbulent boundary layer, the details of which depend on climatic conditions, local topology, and the height above the Earth's surface. The forces are random in nature and, therefore, the structural response can only be determined statistically. Thus, mean and extreme wind loading statistics are used for structural design and form the basis of the wind engineering aspects of building codes. *See* BOUNDARY-LAYER FLOW; ENVIRONMENTAL FLUID MECHANICS; TURBULENT FLOW; WIND.

Of course, there are innumerable other practical examples of this type of flow-induced vibration from the turbulent buffeting of heat exchanger tubes to the gust loading of airplane wings. However, the approach to engineering analysis is basically the same: Compute the structural response statistics using measured flow excitation data. There will always be some turbulence response, but proper design can keep this within acceptable limits. Usually, structural failure due to turbulent buffeting takes a long time to develop and appears in the form of high-cycle fatigue or wear in the case of the rubbing or impacting of loose components. *See* PLASTIC DEFORMATION OF METAL.

Structural failures due to extreme loading caused by such events as hurricane force winds are due to the extremely high average value of the fluid forces and are typically not due to flow-induced vibrations.

Self-excited vibrations. Self-excited vibrations are characterized by periodic fluid excitation forces which depend totally on the structural response; that is, in the absence of structural response, the periodic forces associated with self-excitation mechanisms do not exist. This makes these phenomena much more difficult to study because not only do the structural geometry and fluid mechanics need to be properly scaled but so do the structural dynamics and the fluid-structure interaction. Thus, a model must deform and vibrate just like the full-scale structure (structural dynamic similarity) with the further requirement that the model response in the flow scale to the prototype (fluid-structure similitude). This is often difficult or impossible to achieve fully, and significant judgment may be required for modeling such problems. *See* DIMENSIONAL ANALYSIS.

Mathematically, self-excited vibrations are dynamic stability problems in which the fluid forces are dependent on time and structural displacement. When the flow velocity reaches the critical velocity for instability, V_{cr}, a dramatic increase in oscillation amplitude occurs (Fig. 1, point D). The velocity V_{cr} usually represents the practical upper limit for the flow velocity to prevent structural failure.

Bending-torsion flutter of aircraft wings is an excellent example of self-excited vibrations. As the flow velocity increases, the flow over the wings couples with the structural natural frequencies and modifies them. At the critical velocity for flutter, the aerodynamic coupling causes coalescence of the bending and torsional wing vibration modes and the oscillation amplitudes grow extremely rapidly until catastrophic failure occurs. This happens without warning and so quickly that there is no time for pilot reaction. Thus, bending-torsion flutter is a critical design criterion for fixed-wing aircraft. Interestingly, modern feedback control technologies are used on tactical aircraft to sense impending flutter, and automatically adjust controls to effectively delay the critical velocity. *See* CONTROL SYSTEMS.

There are numerous other examples self-excited vibrations that can cause structural failures in service. Tube-and-shell heat exchangers are used widely in chemical, fossil, and nuclear power plants. Since the 1960s, there have been many failures of tubes in such heat exchangers due to a self-excited phenomenon called fluid-elastic instability. Above the critical velocity, the tubes vibrate with such violence that they can fail rapidly from fatigue or wear due to

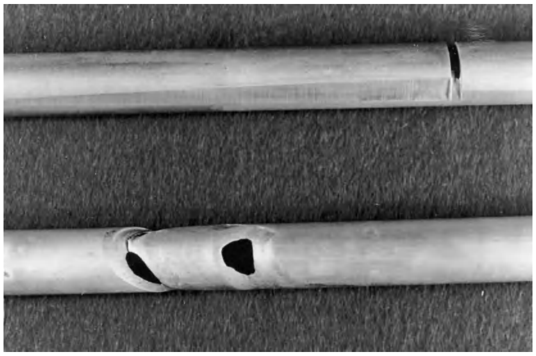


Fig. 2. Heat exchanger tubes cut through at supports and damaged by impacting from broken tubes due to flow-induced vibration.

tube-to-tube clashing. This causes plant shutdowns, loss of productivity, and very costly repairs. In the case of nuclear steam generators it causes leakage of irradiated primary side coolants. Clearly, fluid-elastic instability must be avoided in such equipment to prevent serious damage (**Fig. 2**). *See* HEAT EXCHANGER.

Other examples of self-excited vibrations are seen in the so-called galloping of bluff bodies, such as bridge decks, certain structural shapes, and iced power transmission lines, and the destructive vibrations of hydraulic gates, such as are used for controlling flows in waterways, hydroelectric power plant, and so forth. While the structural geometries and fluid flow details in all these cases are very different, the underlying excitation mechanism is basically the same. The flexibility of the structure permits displacements due to the flowing fluid which fundamentally alters the fluid forces. If these structural-displacement-dependent fluid forces act in such a way as to increase the structural displacement, then the possibility of instability exists. Dynamic instability occurs when the energy transferred from the flow to the structure overcomes the structural damping. Increasing structural stiffness or modifying the structural shape are usually effective means of increasing the maximum permissible flow velocity for self-excited vibrations. Both methods have been used extensively to stabilize the galloping of bridge decks.

Self-controlled or resonant vibrations. Self-controlled vibrations occur when the fluid forces contain strong periodicities. As long as these periodicities are well removed from the structural natural frequencies, the structural response is small, there is no effect of the structural motion on the fluid forces, and the problem can be treated as a forced vibration. However, if the flow periodicity coincides with a structural natural frequency, a large-amplitude resonant response can occur. In this case, the structural response interacts with the flow excitation, increasing its strength and controlling its frequency; thus the appellation “self-controlled.” This response is seen as a sharp peak in the amplitude response (Fig. 1, *B–C*). Resonance can be distinguished from a self-excited instability by the fact that the resonance response disappears at higher flow velocities. Instability responses continue to develop more violently with velocity increases above the stability threshold. *See* RESONANCE (ACOUSTICS AND MECHANICS).

The most common example of the self-controlled excitation mechanism is vortex-induced vibrations of cylinders in a cross flow. As the flow passes the cylinder, vortices are shed alternately into the wake. The flow pattern in the wake is called a von Kármán vortex street and the vortex shedding frequency, f_v, is linearly proportional to the flow velocity, V (**Fig. 3**, *A–E*). This process is scaled using the dimensionless parameter called the Strouhal number, St, where the characteristic length used is the cylinder diameter D. Thus, $\mathrm{St} = f_v D/V$, and has a value of about 0.2 for isolated circular cylinders. *See* DIMENSIONLESS GROUPS; KÁRMÁN VORTEX STREET; VORTEX; WAKE FLOW.

The antisymmetry of the wake vortices produces fluctuating lift forces on the cylinder at the frequency of the shedding of each pair of vortices. Thus, the cylinder vibrations occur primarily in the direction transverse to the flow direction. As the flow velocity is increased, the vortex shedding frequency, f_v, approaches the structural natural frequency, f_n. At frequency coincidence, $f_v = f_n$, resonance occurs (Fig. 3 point *B*). As the vibration amplitudes build up, the fluid lift forces increase in strength and the structural motion dominates the fluid mechanics, locking the vortex shedding frequency onto the structural natural frequency (Fig. 3, *B–C*). Ultimately, the lock-in is lost, the vibration amplitudes drop precipitously, and the vortex shedding frequency jumps back to the natural shedding frequency (Fig. 3, *C–D*). As the flow velocity is further increased, resonance in higher modes may occur.

Probably the most common example of the self-controlled mechanism is the vortex-induced vibrations of smoke stacks in a cross wind. However, other examples abound from submarine periscopes and marine structures in tidal currents to thermocouple probes in pipelines and bluff structures in a wind such as bridge decks. While vortex-induced vibrations in air flows rarely produce structural failures, it is common to use devices such as spiral strakes on smokestacks to minimize the vibrations. On the other hand, vortex shedding in liquid flows can cause serious structural (fatigue) failures so that design of such structures must consider these effects.

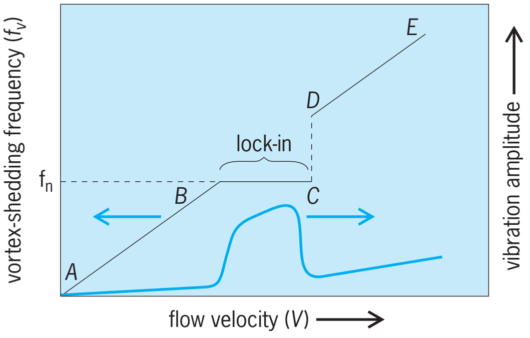


Fig. 3. Vortex-shedding frequency and vibration-response amplitude plotted against flow velocity, showing resonance and frequency control (lock-in).

Some structures are subject to both vortex shedding and self-excited vibrations. For example, bluff bodies such as bridge decks may be excited by both vortex shedding and galloping, and heat exchanger tube arrays experience both vortex shedding and fluid-elastic instability. In some cases, a large-amplitude vortex-shedding resonance can actually precipitate the onset of self-excited vibrations. In other cases, these phenomena occur at quite different flow velocities and behave independently.

David S. Weaver

Bibliography. M. K. Au-Yang, *Flow-Induced Vibration of Power and Process Plant Components*, ASME Press, 2001; R. D. Blevins, *Flow-Induced Vibration*, 2d ed., Van-Nostrand Reinhold, 1990, reprint, Krieger, 2001; E. Naudascher and D. Rockwell, *Flow-Induced Vibrations: An Engineering Guide*, Dover, reprint, 2005; M. P. Païdoussis, *Fluid-Structure Interaction: Slender Structures and Axial Flow*, vol. 1, Academic Press, London, 1998, vol. 2, Elsevier Academic Press, London, 2004; E. Simiu and R. Scanlan, *Wind Effects on Structures: Fundamentals and Applications to Design*, 3d ed., Wiley, 1996.

Flow measurement

The determination of the quantity of a fluid, either a liquid, vapor, or gas, that passes through a pipe, duct, or open channel. Flow may be expressed as a rate of volumetric flow (such as liters per second, gallons per minute, cubic meters per second, cubic feet per minute), mass rate of flow (such as kilograms per second, pounds per hour), or in terms of a total volume or mass flow (integrated rate of flow for a given period of time).

Flow measurement, though centuries old, has become a science in the industrial age. This is because of the need for controlled process flows, stricter accounting methods, and more efficient operations, and because of the realization that most heating, cooling, and materials transport in the process industries is in the form of fluids, the flow rates of which are simple and convenient to control with valves or variable speed pumps. *See* PROCESS CONTROL.

Measurement is accomplished by a variety of means, depending upon the quantities, flow rates, and types of fluids involved. Many industrial process flow measurements consist of a combination of two devices: a primary device that is placed in intimate contact with the fluid and generates a signal, and a secondary device that translates this signal into a motion or a secondary signal for indicating, recording, controlling, or totalizing the flow. Other devices indicate or totalize the flow directly through the interaction of the flowing fluid and the measuring device that is placed directly or indirectly in contact with the fluid stream. **Table 1** shows various metering devices, their operating principles, and flow ranges.

Volume Flow Rate Meters

Flow rate meters may modify the flow pattern by insertion of some obstruction or by modification of the shape of the walls of the conduit or channel carrying the fluid or they may measure the velocity without affecting the flow pattern.

Differential-producing primary devices. Differential-producing primary devices (sometimes called head meters) produce a difference in pressure caused by a modification of the flow pattern. The pressure difference is based on the laws of conservation of energy and conservation of mass. The law of conservation of energy states that the total energy at any given point in the stream is equal to the total energy at a second point in the stream, neglecting

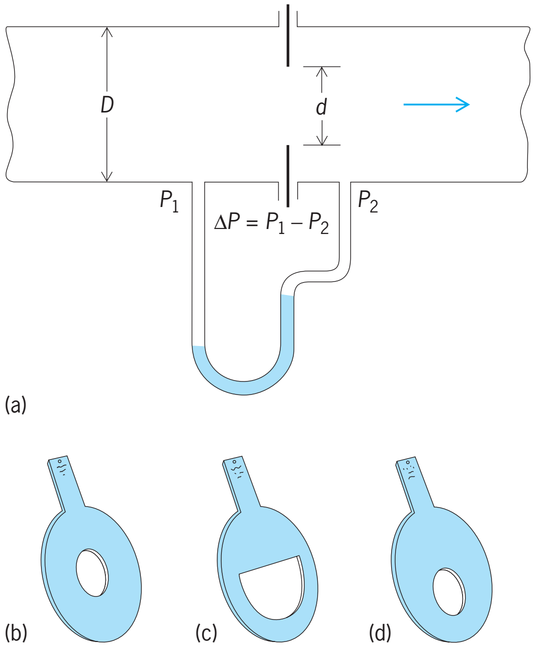


Fig. 1. Orifice plate primary device. (*a*) Schematic cross section. (*b*) Concentric plate. (*c*) Segmental plate. (*d*) Eccentric plate.

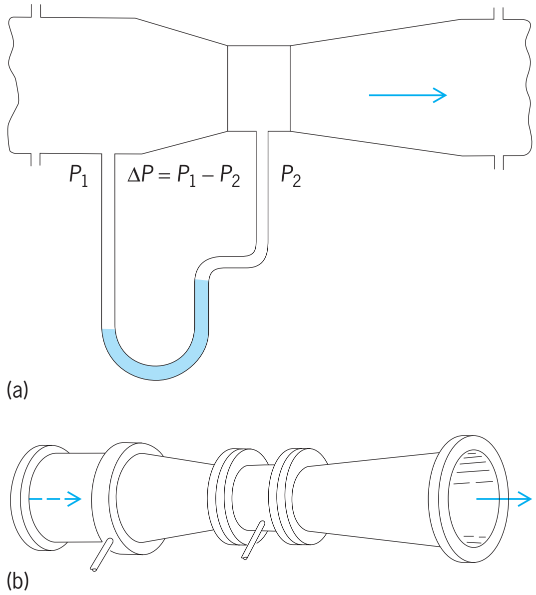


Fig. 2. Venturi tube. (*a*) Schematic cross section. (*b*) Perspective view.

TABLE 1. Examples of fluid meters

Name or type	Operating principle	Useful flow range
Orifice	Differential pressure (ΔP) over a restriction in the flow	3.5:1
Venturi		4:1
Nozzle		4:1
Flow tubes		4:1
Pitot-static	Compares impact pressure with the static pressure	3.5:1
Averaging pitot		3.5:1
Pipe elbow	Pressure from change in it	3.5:1
Target	Force on circular obstruction	3.5:1
Weir	Variable head, variable area	75:1
Flume	Variable head, variable area	up to 100:1
Rotameter (free float)	Automatic variable area	10:1
Piston		10:1
Plug		10:1
Magnetic	Electromagnetic induction	20:1
Tracere	Transit time	10:1
Nutating disk	Positive displacement; separation into and counting of discrete fluid quantities (most of these are available with high accuracy for custody transfer, with a less expensive, less accurate one for monitoring uses)	20:1
Rotary vane		25:1
Oval-shaped gear		50:1
Lobed impeller		10:1
Abutment rotor		25:1
Drum gas meter		40:1
		10:1 to 25:1
Gas turbine meter	Converts velocity to proportional rotational speed	10:1 to 25:1
Liquid turbine		10:1 to 50:1
Bearingless		10:1
Insertion turbine		25:1
Propeller		15:1
Helical impeller		15:1
Vortex cage		15:1
Axial flow mass	Momentum effect due to change in direction of fluid path	10:1
Gyroscopic/Coriolis		>10:1
Gyroscopic		10:1
Thermal loss	Heat extracted by fluid	10:1
Vortex flowmeter	Counting vortices shedding from an obstruction	8:1 to 100:1
Ultrasonic detection of vortex		8:1 to 100:1
Swirl	Precessing rotating vortex	10:1 to 100:1
Fluidic oscillator	Coanda wall attachment	30:1
Ultrasonic contrapropagating: clamp-on built-in spool piece	Time difference upstream/downstream	25:1
Ultrasonic contrapropagating: multiple path in spool piece		
Ultrasonic contrapropagating: axial in spool piece		
Ultrasonic correlation	Transit time of turbulence	10:1
Ultrasonic correlation: deflection (drift)	Deflection of beam across pipe	5:1
Ultrasonic contrapropagating: Doppler	Doppler shift-reflection from particles	10:1
Noise measurement	Spectral distribution of noise	5:1
Nuclear magnetic resonance	Transit time of nuclear disturbance	10:1
Laser Doppler velocimeter	Frequency shift of light reflected from particles	>1000:1

frictional and turbulent losses between the points. It is possible, however, to convert pressure (potential energy) to a velocity (kinetic energy), and vice versa. By use of a restriction in the pipe, such as an orifice plate (**Fig. 1**), venturi tube (**Fig. 2**), flow nozzle (**Fig. 3**), or flow tube (**Fig. 4**), a portion of the potential energy of the stream is temporarily converted to kinetic energy as the fluid speeds up to pass through these primary devices. *See* BERNOULLI'S THEOREM.

The secondary device measures this change in energy as a differential pressure ΔP, which is related to the flow by Eq. (1), where V is the velocity of flow

$$V = K_1\sqrt{\Delta P/\rho_f} \tag{1}$$

through the restriction and ρ_f is the mass density of the flowing fluid. The constant K_1 includes the orifice or throat diameter and the necessary dimension units, sizes, local gravity, and so forth. It also contains the discharge coefficient which has been experimentally determined and reduced to equation form for certain standard differential producers [namely, the square-edged orifice with specified pressure taps, ASME (American Society of Mechanical Engineers) nozzles, and classical venturis].

Mass flow for liquids, q_m, is given by Eq. (2), where K_2 is given by Eq. (3), and the quantities in Eqs. (2)

$$q_m = K_2\sqrt{\Delta P \rho_f} \tag{2}$$

$$K_2 = NSD^2 \tag{3}$$

and (3) are given in **Table 2**. The approximate value

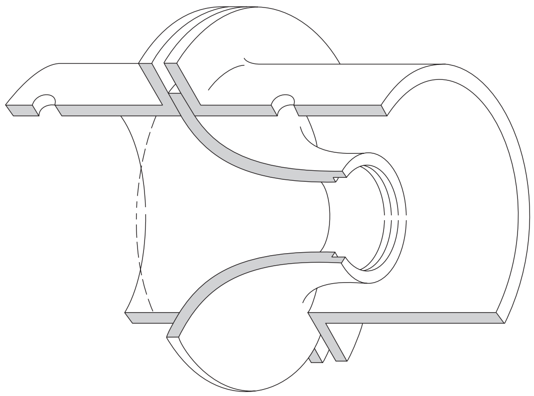


Fig. 3. Cutaway view of a flow nozzle.

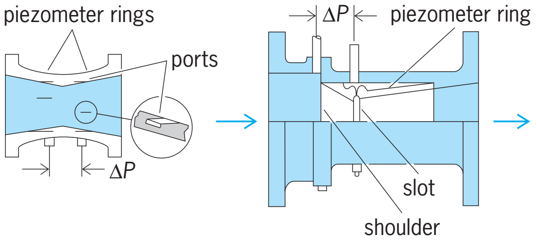


Fig. 4. Two types of flow tube. (*After 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*)

of the sizing factor S, upon which K_2 depends, can be found from **Fig. 5**.

Beyond the restriction, the fluid particles slow down until the velocity is once more the same as that in the upstream conduit section. However, the static pressure is less due to the effect of turbulent and frictional losses associated with the acceleration and deceleration at the restriction. The magnitude of the loss depends upon the configuration of the primary device. *See* METERING ORIFICE.

Orifice plate. The common orifice plate (Fig. 1) is a thin plate inserted between pipe flanges, usually having a round, concentric hole with a sharp, square upstream edge. The extensive empirical data available and its simplicity make it the most common of the primary devices. Segmental and eccentric orifices are useful for measurement of liquids with solids, and of vapors or gases with entrained liquids.

Venturi tube. This device (Fig. 2) has the advantage of introducing less permanent loss of pressure than the orifice plate. The converging inlet cone permits solids and dirt to be flushed through it, and the outlet cone reduces the turbulent losses.

Flow nozzle. The flow nozzle (Fig. 3) also has a converging inlet, but it has no diverging outlet section. The pressure loss is similar to that with an orifice plate for the same differential pressure. It is used where solids are entrained in the liquid and also where fluid velocities are particularly high (it will pass more fluid than an orifice of the same diameter).

Flow tubes. Several designs of primary devices have been developed to give lower permanent pressure loss than orifices and shorter overall length than standard venturis, and they sometimes incorporate piezometer chamber averaging of the pressure tap. Figure 4 shows two such designs.

Pitot tubes. In industrial flow measurement, pitot tubes are used in spot checks and for comparative measurements to determine trends. A pitot tube measures the difference between total impact pressure and static pressure. In the device shown in **Fig. 6** the central tube receives the impact pressure, while the holes in the outer tube are subjected to the static pressure. A secondary device senses the difference in these pressures and converts it into an indication of fluid velocity at the location of the tip. There are a wide variety of other tube configurations available for sensing both impact and static pressure or combinations thereof. Several readings should be taken across the pipe along a diameter (preferably two diameters) of the pipe and the average velocity determined. The pitot tube is used extensively as a laboratory device for the point velocity measurement and also for speed measurement of vehicles and in wind tunnel work.

The averaging pitot tube stimulates an array of pitot tubes across the pipe. It consists of a rod extending across the pipe with a number of interconnected upstream holes and a downstream hole for the static pressure reference. In one design, the diamond shape of the rod is claimed to present a very

TABLE 2. Quantities used in liquid-flow calculation

Symbol	Quantity	Unit or value		
		SI units	U.S. Customary units	
N	Constant that depends upon the system of units used	1.11	0.0999	0.52
D	Inside diameter of pipe	meter	inch	inch
ΔP	Pressure difference	pascal	in. H_2O*	lbf/in.^{2} (psi)
ρ_f	Density of the flowing fluid	kg/m^3	lb$_m$/ft^3	lb$_m$/ft^3
β	Ratio, (d/D)	dimensionless		
C	Discharge coefficient	dimensionless		
S	Sizing factor $= C\beta^2/\sqrt{1-\beta^4}$	dimensionless		
d	Orifice or throat diameter	meter	inch	inch
q_m	Mass flow rate	kg/s	lb$_m$/s	lb$_m$/s

*The pressure of a column of water 1 in. (25.4 mm) high at 68 °F (20 °C) and acceleration of gravity of 32.174 ft/s^2 (9.80665 m/s^2).

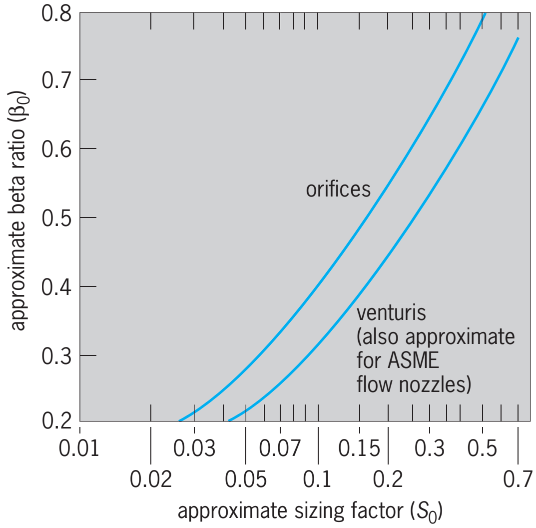


Fig. 5. Sizing factors for differential-primary-producing elements, for liquid flows, based on $C = 0.6 + 0.06\ S$ for orifices, $C = 0.995$ for venturis. The factors do not include the Reynolds number correction. (*Foxboro Co.*)

low flow obstruction and still, with the interconnected ports, to provide a differential pressure that is representative of the flow relatively unaffected by profile differences. *See* PITOT TUBE.

Pipe elbow. Fluid flow around the bend in a pipe elbow (**Fig. 7**) produces a centrifugal force greater on the outside than on the inside, and thereby produces a difference in pressure between pressure connections on the inside and outside of the bend. This differential pressure is used as an indication of the volume flow rate. This device is often used for rough flow measurement because it is not necessary to disturb existing piping. Carefully installed and calibrated, the device can provide an accurate measurement.

Differential-pressure secondary devices. All the primary devices discussed above create a difference in pressure between two points. This differential pressure must be measured and converted by the equations given to obtain flow. Differential pressure may be measured accurately by simple liquid-filled (usually mercury) U-tube manometers or by more refined types of meters, such as the float-operated mechanism or the weight-balanced, ring-type meter, to provide controlling, recording, and totalizing functions. These devices (**Fig. 8**) are ordinarily connected directly to the pressure connections on the primary device and are therefore exposed to the process fluids. Sometimes an inert-liquid-seal fluid is used to isolate the secondary device from corrosive process fluids. In either case, the manometers are best located adjacent to the primary device. These manometers are accurate if the mercury is clean and the fluid flow rate is not changing too rapidly. *See* MANOMETER.

Dry-type sensors. Diaphragm- and bellows-type differential-pressure measuring devices were developed to eliminate manometer fluids and to provide faster response and easier installations. **Figure 9** shows a representative bellows-type differential-pressure sensor. The range of these instruments is changed by removing one cover and adjusting the range spring or replacing it with one of a different spring rate. Thermal expansion of the fill fluid contained in the two bellows is compensated for by the bimetal-supported additional convolutions shown at the left of the figure. These are very stable and reliable instruments.

Many direct-connected meters have been modified to provide a pneumatic or electric signal to remote instrumentation. Serviceable transmitters have been developed. Two types of meters, the force-balance and deflection types, are suited to either pneumatic or electric transmission of the flow signal.

The force-balance type (**Fig. 10**) causes the variation of an air pressure in a small bellows to oppose the net force created by the differential pressure on opposite sides of a diaphragm or diaphragm capsule. This air pressure is used as a signal representing the flow rate. An electronic force-balance design utilizes a restoring force due to a force motor and uses an inductive detector for unbalance; the current in the motor is the transmitted (4–20 mA) signal. One electronic version provides an output signal proportional to flow by making the feedback force the interaction between two motors, with the current flowing through them (the force is squared).

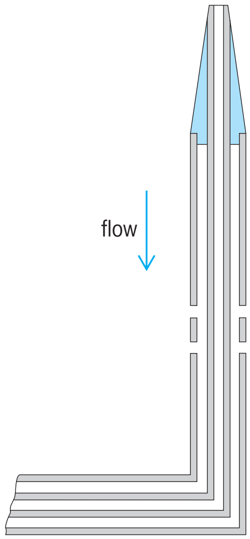


Fig. 6. Laboratory-type pitot tube.

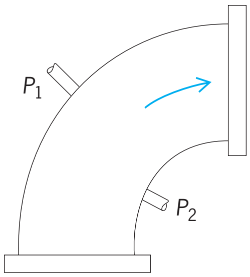


Fig. 7. Pipe elbow. Difference between outside pressure P_1 and inside pressure P_2 indicates volume flow rate.

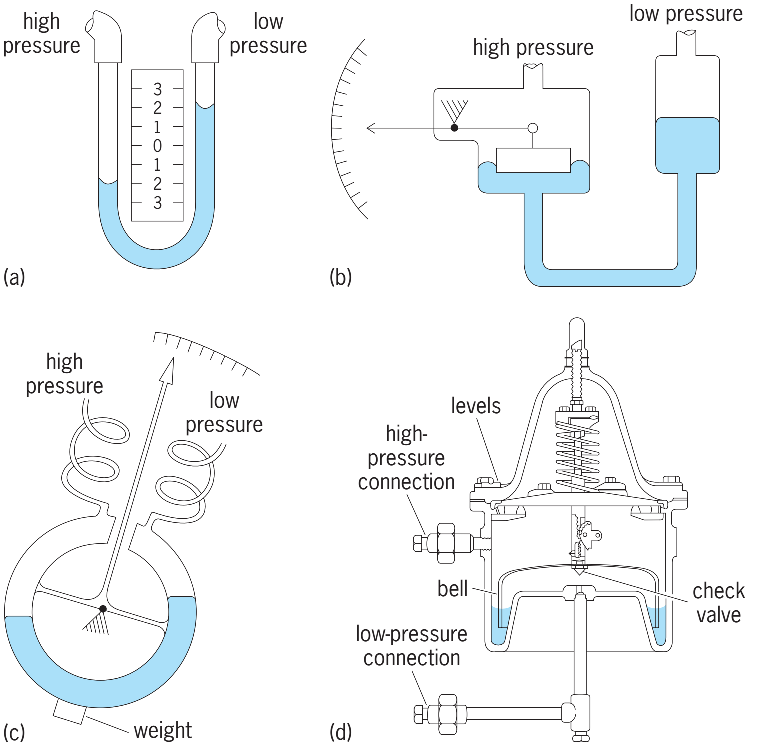


Fig. 8. Differential-pressure secondary devices. (*a*) U-tube manometer. (*b*) Mercury float mechanism. (*c*) Weight-balanced ring-type meter. (*d*) Bell type (*Foxboro Co*.).

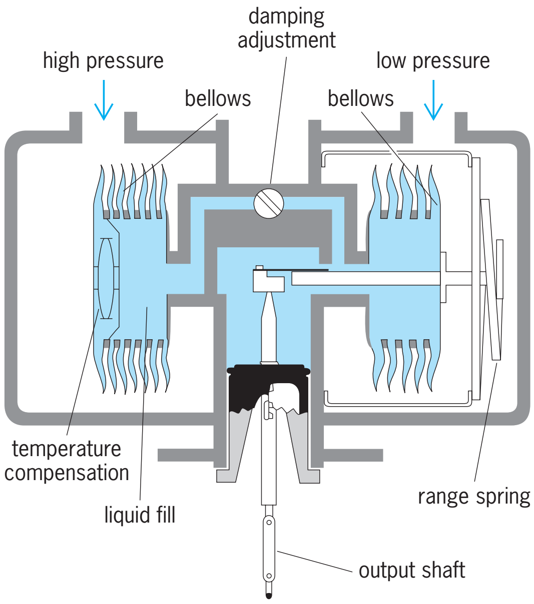


Fig. 9. Bellows-type differential-pressure-measuring element. (*Foxboro Co*.)

Many types of deflection differential-pressure cells are offered commercially. Several apply the differential pressure across a diaphragm containing or connected to a strain gage (bonded-semiconductor, diffused strain gage in a silicon diaphragm, and so forth; **Fig. 11*a***). A representative two-wire circuit is shown in Fig. 11*b*. The signal and power are carried by the same two wires.

Some differential-pressure cells use variable capacity or differential capacity between a diaphragm and an adjacent wall (Fig. 11*c*). Some use the moving diaphragm to cause a variable reluctance or variable inductance. In most cases the measuring elements are protected from the fluid being measured by a pair of seal-out diaphragms or bellows, and the elements are submerged in silicon oil. In one type the measuring diaphragm provides a variable tension on a wire that is placed in a magnetic field and connected to a self-oscillating circuit that provides a frequency that varies with the differential pressure (Fig. 11*d*).

Readout scales. Since these secondary devices measure the differential pressure developed, their outputs are proportional to the square of the flow rate. In the case of direct-reading devices, they must be fitted with square root scales and charts for reading flow rate. These scales are particularly difficult to read at low values, thereby limiting the useful rangeability. Uniform flow-rate scales are obtained with manometers with specially shaped chambers or with shaped plugs (Ledoux bell) or with special cams in the linkage between the manometer and the indicating pointers or pens. Some differential-pressure transmitters provide linear-with-flow transmitted signals by incorporating squaring functions in the feedback, or mechanical or electronic square rooting can be done either at the transmitter or at the receiver location.

Target meter. The target-type meter (**Fig. 12**), besides being particularly suitable for dirty fluids, has the important advantage of a built-in secondary device and is without the disadvantage of fluid-filled connection pipes to the secondary device.

Open-channel meters. Flow in open channels is determined by measuring the height of liquid passing through a weir or flume. Consistent results require that the measuring device be protected from direct impact of the flowing stream. Because the height-flow velocity relationship is exponential, the secondary device frequently uses cams to obtain a uniform flow scale.

Variable-area meters. These meters work on the principle of a variable restrictor in the flowing stream

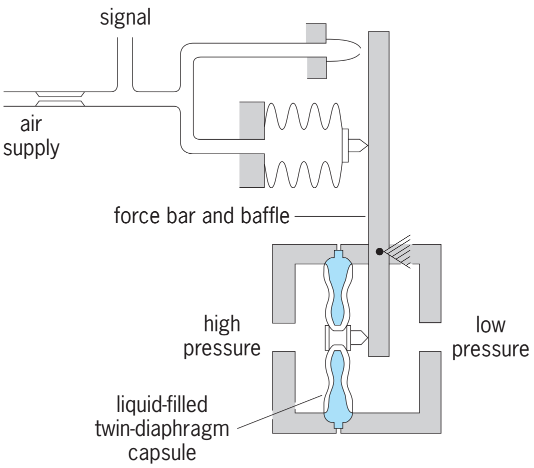


Fig. 10. Diaphragm meter, force balance type.

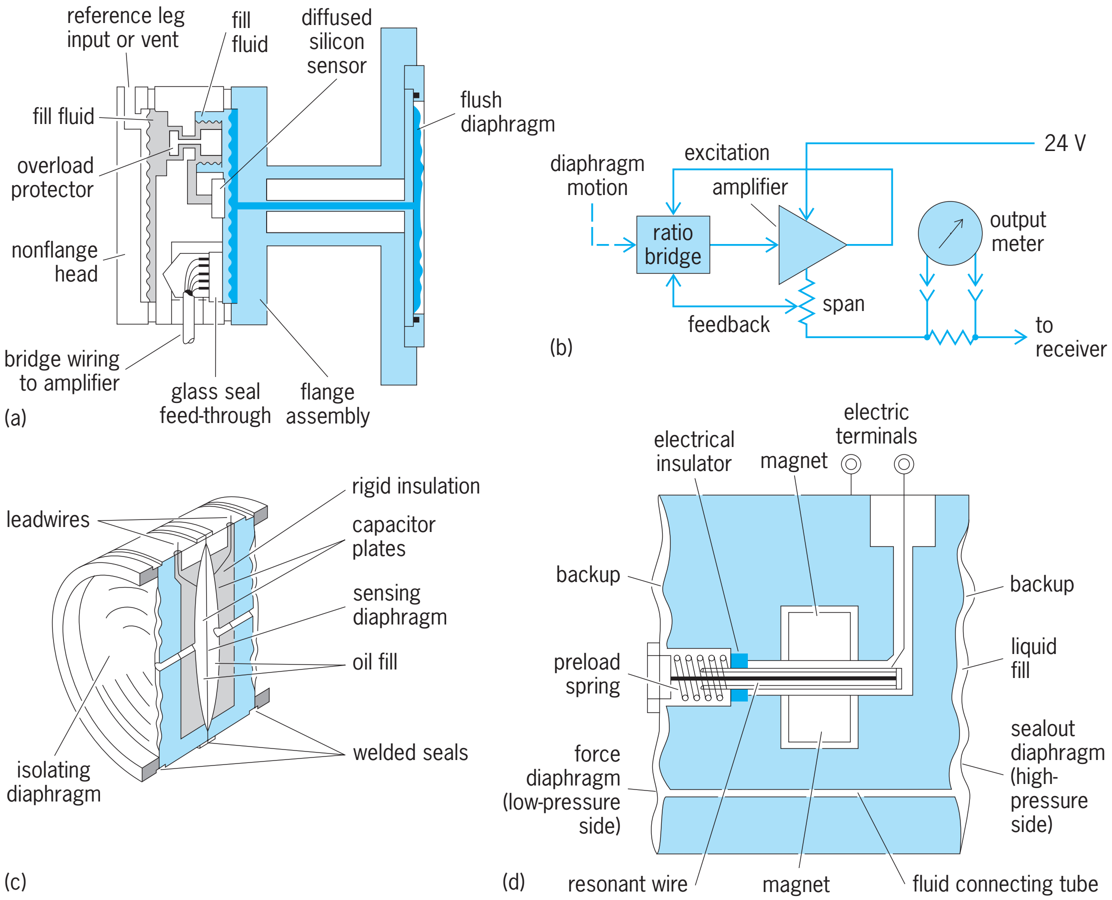


Fig. 11. Differential-pressure sensors. (*a*) Silicon-chip type (*Honeywell Industrial Division*) with (*b*) 4- to 20-mA two-wire transmitter circuit (*after 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*). (*c*) Capacitance type (*Rosemount Engineering Corp.*). (*d*) Resonant-wire type (*Foxboro Co.*).

being forced by the fluid to a position to allow the required flow-through. Those that depend upon gravity provide essentially a constant differential pressure, while those that supplement the gravity with a spring will have a variable differential pressure.

In the tapered-tube rotameter (**Fig. 13*a***), the fluid flows upward through a tapered tube, lifting a shaped weight (possibly misnamed a float) to a position where the upward fluid force just balances the float weight, thereby giving a float position that indicates the flow rate. For remote indication, and also where the pressure requires a metal tube, a center rod attached to the float carries the core of an inductance pickup for transmission of the float position through a nonmagnetic tube seal. In other designs the rod carries one side of a magnetic coupling to a follower mechanism outside of the nonmagnetic tube.

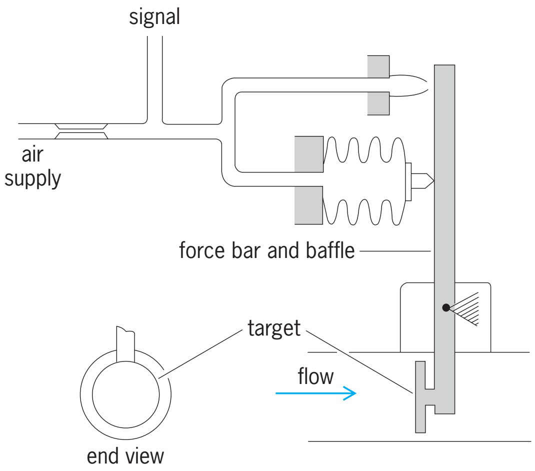


Fig. 12. Target flowmeter. (*Foxboro Co.*)

In a position type (Fig. 13*b*) the bouyant force of the liquid carries the piston upward until a sufficient area has been uncovered in a slot in the side of the vertical tube to allow the liquid to flow through the slot. The position of the piston indicates the flow rate and, by properly shaping the slot, this type of meter may be made with a uniform flow scale. A spring is used in some designs to supplement the gravity of the plug and further characterize the indication.

In the tapered-plug design (Fig. 13*c*), the plug, located in an orifice, is raised until the opening is sufficient to handle the fluid flow. Readouts can be the same as in the tapered-tube design.

A spring-loaded variable-area meter (Fig. 13*d*) provides a differential pressure for a secondary

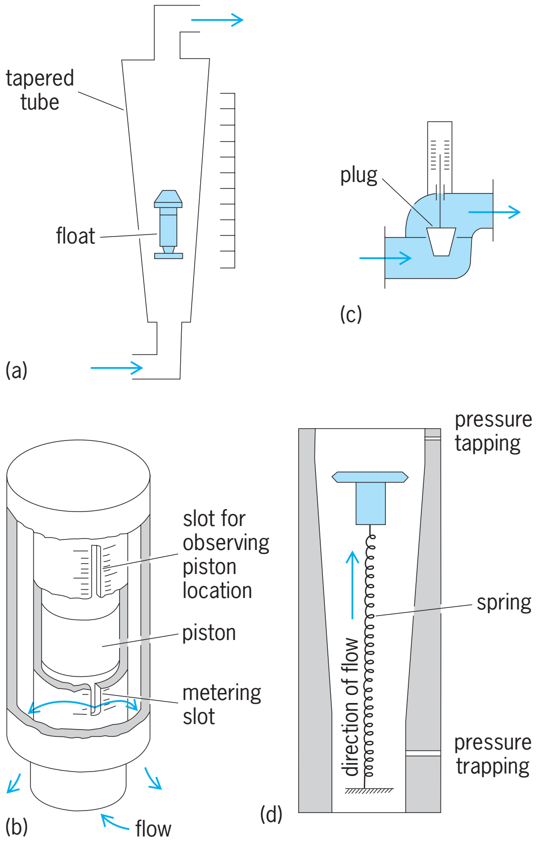


Fig. 13. Variable-area flowmeters. (*a*) Rotameter type. (*b*) Piston type. (*c*) Plug type (*after 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*). (*d*) Spring-loaded differential-pressure producer (*after Institute of Mechanical Engineers*, *How to Choose a Flowmeter*, *1975*).

device. By combining the characteristics of a spring-loaded variable-area meter and a standard differential-pressure measurement across the float, it is claimed that the rangeability of the measurement can be raised to as high as 50:1.

Electromagnetic flowmeter. A magnetic field is applied across a metering tube carrying a flowing liquid, and the voltage generated between two perpendicularly located electrodes contacting the liquid is measured by a secondary or transmitter instrument. Since this voltage is strictly proportional to the average velocity, a linear volume flow signal results. Practical flow measurements can be made on any liquid that is at all conductive. The linear scale allows accurate flow measurements over a greater range than is possible in the differential-pressure meters. The lack of obstruction to fluid flow makes it possible to measure thick slurries and gummy liquids. In order to avoid errors due to electrodes becoming coated with an insulating material, some electromagnetic flowmeters are equipped with ultrasonic electrode cleaners. *See* ELECTROMAGNETIC INDUCTION.

Classical electromagnetic flow tubes power the field coils with a continuous ac mains voltage and measure the induced ac voltage on the electrodes. After an approximately 90° phase shift, a ratio is made of the two voltages to correct for mains voltage change. Some electromagnetic flowmeters apply a periodic dc voltage pulse to the coils and measure the dc voltage across the electrodes at a carefully selected time relative to the field excitation. These meters are available with nonmagnetic stainless steel pipe and various types of insulating lining, and with all-plastic construction. When the field coils are mounted inside the pipe, a magnetic iron may be used for the pipe.

Tracer method. This classic method consists of injecting a small quantity of salt into the flowing stream and measuring the transit time of passage between two fixed points by measuring electrical conductivity at those points. Inaccuracies are introduced by variations in flow cross section in the pipe or channel being measured. Because of the convenience for a quick check of flow under comparatively adverse conditions, standards have been written around the use of salt, radioactive material, and even pulses of heat. Appropriate sensors are offered with correlation techniques for improving the accuracy.

Quantity Meters

Quantity meters are those that measure or pass fixed amounts of fluid. By introducing a timing element, all quantity meters can be converted to volume flow-rate meters. Quantity meters may be classified as positive-displacement meters or rotating impeller-type meters. These are illustrated in **Figs. 14** and **15**.

Positive-displacement meters. The flowing stream is separated into discrete quantities by these meters, which capture definite volumes one after another and pass them downstream. Some representative examples will be discussed.

Nutating-disk meter. A circular disk attached to a spherical center is restrained from rotation by a vertical partition through a slot in the disk (Fig. 14*a*). It is allowed to nutate so that its shaft is in contact with a conical guide. The liquid entering the inlet port moves the disk in a nutating motion until the liquid discharges from the outlet port. The double conical shape of the measuring chamber makes a seal with the disk as it goes through the nutating motion. Normally a mechanical counter indicates the number of cycles as the fluid flows through smoothly with no pulsations.

Rotary-vane meter. An eccentrically mounted drum carries radial spring-loaded vanes (Fig. 14*b*). These vanes slide in and out, and seal against the meter casing to form pockets that carry a measured amount of fluid on each cycle.

Oval-shaped gear meter. Two carefully formed gears provide their own synchronization, and with close-fitting teeth to the outer circular-shaped chambers they deliver a fixed quantity of liquid for each rotation (Fig. 14*c*). Rotation of the output shaft can be recorded smoothly by a contact on each revolution. If a higher-frequency signal is desired for flow control or closer monitoring, a compensating gear or optical disk can be used to make the output uniform for a constant liquid flow.

Lobed impeller. The lobed impeller (Fig. 14*d*) is similar to the oval gear except that the synchronization is done by external circular gears and therefore the rotation is smooth.

Rotary-abutment meter. Two displacement rotating vanes interleave with cavities on an abutment rotor (Fig. 14*e*). The three elements are geared together, and the meter depends upon close clearances rather than rubbing surfaces to provide the seals.

Liquid-sealed drum-type gas meter. A cylindrical chamber is filled more than half full with water and divided into four rotating compartments formed by trailing vanes (Fig. 14*f*). Gas entering through the center shaft from one compartment to another forces rotation that allows the gas then to exhaust out the top as it is displaced by the water.

Remotely indicating meters may incorporate a simple electronic contact, or a magnetic or optical sensor which can transmit one or more pulses per revolution.

Rotating impeller meters. These convert velocity to a proportional rotational speed. They include various types of turbine meters, propellers, helical impellers, and vortex cage meters.

Turbine meters. The gas turbine meter (Fig. 15*a*) is characterized by a large central hub that increases the velocity of the gas and also puts it through the tips of the rotor to increase the torque. Readout can be by a directly connected mechanical register or by a proximity sensor counting the blade tips as they pass by. Low-friction ball bearings require special lubricants and procedures.

Most of the available types of liquid-turbine meters (Fig. 15*b*) are designed to cause a fluid flow or hydraulic lift to reduce end thrust, and several depend upon the flowing fluid to lubricate the bearings.

One bearingless-turbine meter is designed to float a double-ended turbine on the flowing fluid and thereby avoid the problem of wearing bearings. The turbine rotor can be made of various materials to suit the fluid, making it an attractive flow measurement device for nonlubricating and corrosive fluids.

In one insertion meter a small propeller or turbine rotor is mounted at right angles to the end of a support rod. Readout is normally through an inductive pickup and appropriate electronics for obtaining rate of flow. The propeller can be inserted into a flowing stream or into a closed pipe with a direct readout of the flow rate at the location of the propeller. Proper location can give an average flow rate either locally or remotely.

Propeller, helical impeller, and vortex cage meters. Since the elements in these meters (**Fig. 16**) rotate at a speed that is linear with fluid velocity, the useful rangeability is large and limited mostly by the bearings. For local indication, revolutions can be counted by coupling to a local-mounted counter. Proximity detectors, giving pulses from 1 to as many as 100 pulses per revolution, can be used for transmitting to remote integrators or flow controllers.

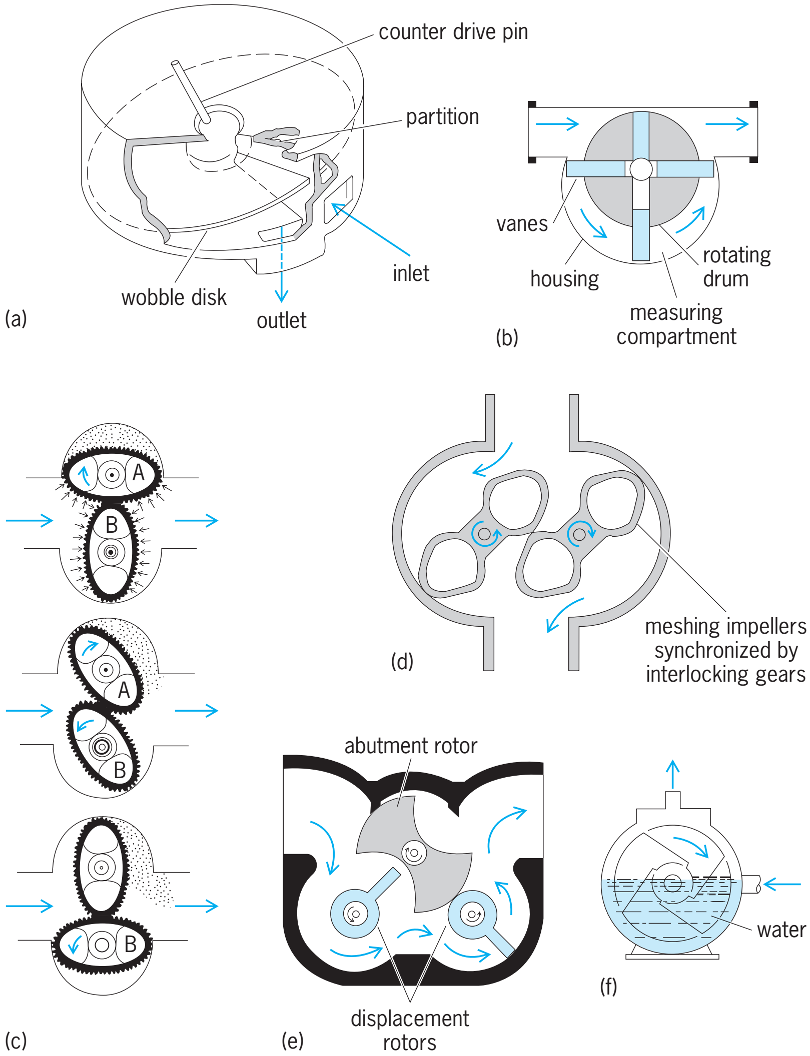


Fig. 14. **Positive-displacement flowmeters. (*a*) Nutating disk. (*b*) Rotary vane. (*c*) Oval-shaped gear (*Brooks Instrument Division*). (*d*) Lobed impeller. (*e*) Rotary abutment. (*f*) Liquid-sealed drum-type gas flowmeter (*after 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*).**

Flow correction factors. The accuracy of quantity meters is dependent to varying degrees on the physical properties of the fluid being measured. Changes from the reference (calibration) conditions of viscosity and density must be considered if accurate measurement is required. It is usually the practice to take pressure and temperature measurements of vapors and gases, or temperature measurement of liquids, at the flow-metering device so that corrections either may be applied by the mechanism of a recording or totalizing device or introduced manually. Some meters are available in which these corrections are introduced automatically with great accuracy, while others are made in which only a first-order approximation is introduced. Proper selection is determined by the requirements of the installation. A meter compensated for fluid conditions can become a mass flowmeter.

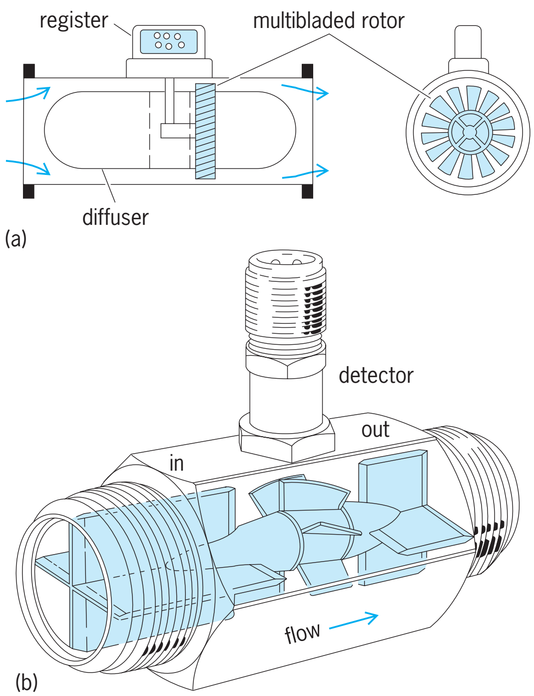


Fig. 15. Turbine flowmeters. (*a*) Gas turbine flowmeter (*after Gas meters, Factory Mag., 9(1):39–45, January 1976*). (*b*) Liquid turbine flowmeter (*Foxboro Co.*).

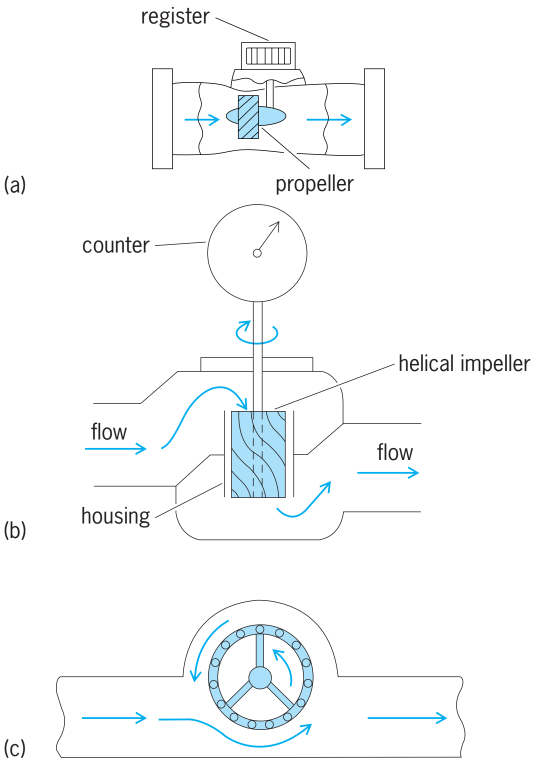


Fig. 16. Types of quantity flowmeters. (*a*) Propeller. (*b*) Helical impeller. (c) Vortex cage.

Mass Flow Rate Meters

There have been many attempts to develop practical industrial mass flowmeters. The most consistently satisfactory method of determining mass flow has been through calculation using volume flow and density determined either by measurement or inference from pressure and temperatures (**Fig. 17**). Only a few of the mechanical type will be mentioned here.

In one type, the fluid in the pipe is made to rotate at a constant speed by a motor-driven impeller. The torque required by a second, stationary impeller to straighten the flow again is a direct measurement of the mass flow (**Fig. 18*a***).

A second type, described as gyroscopic/Coriolis mass flowmeter, employes a C-shaped pipe and a T-shaped leaf spring as opposite legs of a tuning fork (Fig. 18*b*). An electromagnetic forcer excites the tuning fork, thereby subjecting each moving particle within the pipe to a Coriolis-type acceleration. The resulting forces angularly deflect the C-shaped pipe an amount that is inversely proportional to the stiffness of the pipe and proportional to the mass flow rate within the pipe. The angular deflection (twisting) of the C-shaped pipe is optically measured twice during each cycle of the tuning-fork oscillation (oscillating at natural frequency). A digital logic circuit converts the timing of the signals into a mass-flow-rate signal.

In another configuration (using the Coriolis effect) a ribbed disk is fastened by a torque-sensing member within a housing, and both are rotated at constant speed. The fluid enters at the center of the disk and is accelerated radially; the torque on the disk is a direct measure of the mass flow rate. Unfortunately, the practical problem of maintaining the rotating seals in the device has hindered its commercial development. *See* CORIOLIS ACCELERATION.

Another type is based on the principle of the gyroscope (Fig. 18*c*). The flowing stream is sent through a pipe of suitable shape so that the mass of material flowing corresponds to a gyro wheel rotating about axis C. The entire gyrolike assembly is rotating about axis A, and a torque is produced about axis B proportional to the angular momentum of the simulated gyro wheel. This torque is therefore directly proportional to the mass rate of flow. The pretzel-like configuration of pipe is introduced to eliminate centrifugal and other extraneous effects. *See* GYROSCOPE.

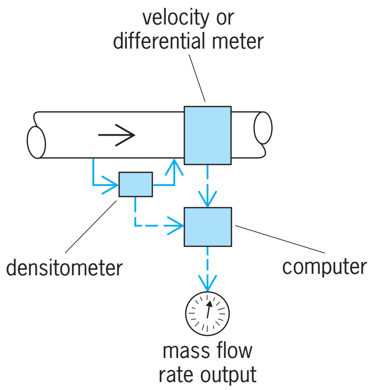


Fig. 17. Inferential mass-flowmeter which computes real-time flow rate in terms of volume rate times gas density. (*After Gas meters, Factory Mag., 9(1):39–45, January 1976*)

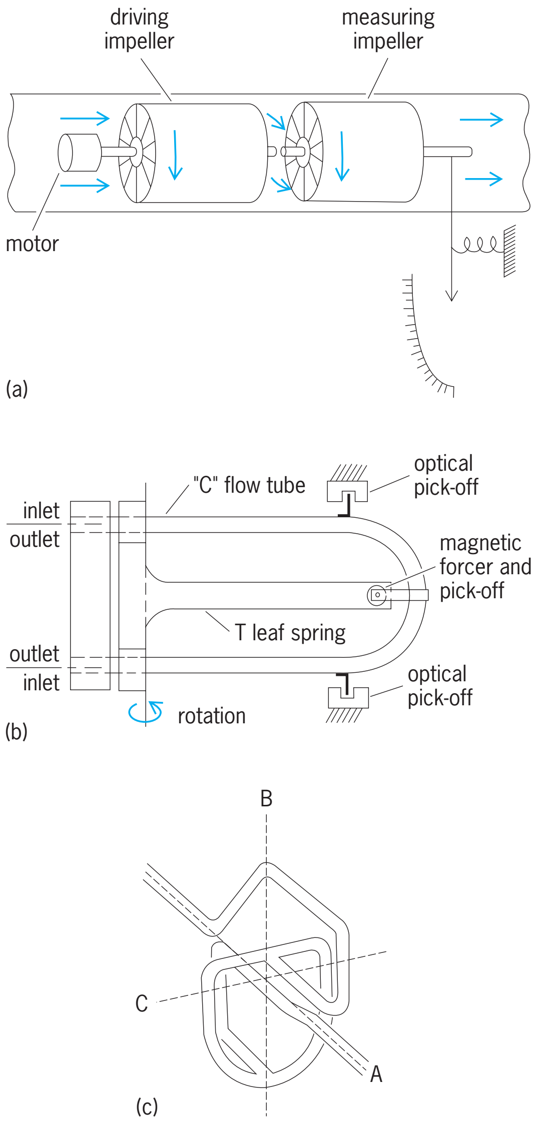


Fig. 18. Mass flowmeters. (*a*) Axial-flow type (*General Electric*). (*b*) Gyroscopic/Coriolis type (*after K. O. Plache, Coriolis/Gyroscopic Flow Meter, ASME Pap. 77-WA/FM/4, American Society of Mechanical Engineers, 1977*). (c) Gyroscopic type.

Thermal Flow Measurement Devices

Several heated or self-heating types of sensor are used for gas-velocity or gas-mass flowmeters.

Hot-wire anemometer. A thermopile heated by a constant ac voltage is cooled by the gas flow. The degree of cooling, proportional to local flow velocity, is indicated by the thermopile millivoltage output. *See* THERMOCOUPLE.

Heat-loss flowmeter. A constant current is supplied to a thermistor bead. The temperature of the bead depends upon the cooling effect of the mass flow of the fluid, and the measure of the voltage drop across the bead is therefore an indication of the flow. A second thermistor bead is used to compensate for the temperature of the fluid. *See* THERMISTOR.

Another heat-loss flowmeter measures the amount of electrical heating that is required to keep the temperature of a platinum resistor constant since the amount of heat carried away changes with the velocity of the fluid. A second sensor is used to compensate for the temperature of the fluid.

Vortex Flowmeters

Flowmeters have been developed based on the phenomena of vortex shedding and vortex precession. *See* VORTEX.

Vortex-shedding flowmeters. One of the newest successful methods of measuring flow is based on one of the oldest observed phenomena relating to flow— the eddies (vortices) forming behind an obstruction in a flowing stream. Close observation has shown that each vortex builds up and breaks loose, followed by a repeat in the opposite rotation on the other side of the obstruction, and so alternately the vortices shed and go downstream in what is called a vortex street. Further observation shows that the distance between successive vortices is directly proportional to the blocking width of the obstruction. The distance is also dependent of the fluid velocity over a large range, and thus the frequency of the shedding is directly proportional to the velocity of the fluid. *See* KÁRMÁN VORTEX STREET.

Application of vortex-shedding flowmeters is normally limited to the turbulent-flow regime, and various shedding-element shapes and internal constructions have been developed in order to improve the shedding characteristics and the detection of the shedding. The common characteristic is sharp edges around which the vortices can form.

Sensing of the shedding vortices is done by (1) measuring the cyclical change in flow across the front surface of the shedder (self-heated resistance elements); (2) sensing the change in pressure on the sides of the "tail" of the shedder (inductive sensing of motion of a "shuttlecock" in a cross passage, or piezoelectric or capacitive sensing of the pressure on diaphragms sealing a cross passage); (3) sensing the forces on the shedder or an extended tail (piezoelectric sensing of the bending of the shedder, or strain-gage sensing of tail motion); (4) sensing of the vortices after they have left the shedder (by an ultrasonic beam across the wake); (5) sensors mounted outside of the pipe at the junction between two tubes connected to openings in the sides of the tail of the shedder.

One design, which uses a rod or a taut wire as the obstruction and as the ultrasonic detector, is offered for open-channel measurement as well as ship velocity.

Vortex flowmeters are available for measuring the flow of liquids, vapors, and gases. The pulse rate output is particularly convenient for input to integrators for recording total flow and for digital computers. The linear-with-flow output makes the measurement particularly attractive for widely varying flows. However, the relatively low frequency (as low as 1 Hz) of the large blockage designs must be considered when using this measurement in a flow-control loop.

Vortex-precession flowmeter. In this instrument, used to measure gas flows, a swirl is imparted to the flowing fluid by a fixed set of radial vanes. As this swirl goes into an expanding tube, it forms a precessing vortex that is detected by a heated thermistor sensor. The frequency of the precession is claimed to be proportional to the volumetric flow rate within an accuracy of 0.75%.

Fluidic-Flow Measurement

Two types of flowmeters based on fluidic phenomena have been developed: the fluidic oscillator meter and the fluidic flow sensor.

Fluidic-oscillator meter. This type of meter works on the principle of the Coanda effect, the tendency of fluid coming out of a jet to follow a wall contour. The fluid entering the meter will attach to one of two opposing diverging side walls (**Fig. 19***a*). A small portion of the stream is split off and channeled back through the feedback passage on that side to force the incoming stream to attach to the other side wall (Fig. 19*b*). The frequency of the oscillation back and forth is directly proportional to the volume flow through the meter. A sensor detects the oscillations and transmits the signal. This principle is used in both liquid and gas flowmeters.

Fluidic-flow sensor. **Figure 20** shows this device (also known as a deflected-jet fluidic flowmeter) as used in gas flows, particularly in dirty environments. It consists of a jet of air or other selected gas directed back from the outer nozzle onto two adjacent small openings. The flow of the gas being measured will deflect the jet to change the relative pressure on the two ports and thereby give a signal corresponding to the gas velocity.

Ultrasonic Flow Measurement

Ultrasonic flowmeters are available in many physical arrangements and are based on a great variety of fundamental principles. Most are found in one of the following categories.

Contrapropagating meter. Contrapropagating diagonal-beam designs (**Fig. 21***a*) have transducers mounted in sockets machined in the pipe or welded on wedges, and also as clamp-on elements. The difference between upstream and downstream "flight" times combined with the speed of sound in the fluid will give fluid velocity. Since the flow near the middle of the pipe is faster, a multiple parallel-beam design (Fig. 21*b*) will give a better average. In a symmetrical profile a beam offset from the center of the pipe may also give a better average.

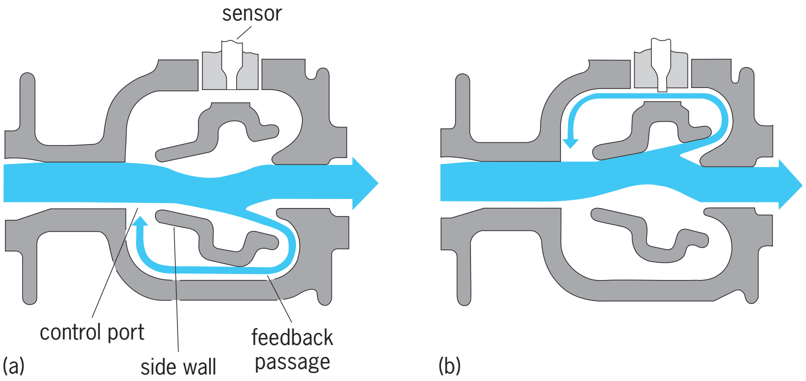


Fig. 19. Fluidic oscillator meter in which feedback stream attaches to (*a*) first one wall, and (*b*) then the other. (*After S. J. Bailey, Tradeoffs complicate decisions in selecting flowmeters, Control E Eng., 27(4):75–79. 1980*)

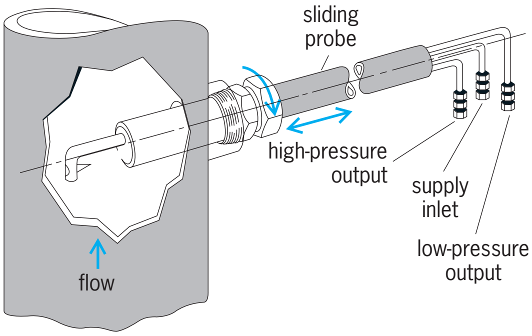


Fig. 20. Fluidic flow sensor. (*Fluid Dynamics Devices*)

Axial meter. For small pipes, the short path length across the fluid stream makes measurement difficult. Figure 21*c* shows a contrapropagating arrangement for measuring lengthwise over a relatively longer path where timing measurements are more practical.

Correlation meter. Two pairs of elements with beams across the pipe can be used with a correlation circuit to detect the time it requires for discontinuities in the fluid stream to pass between the two detectors (Fig. 21*d*).

Deflection (drift) meter. An ultrasonic beam directed across the pipe will be deflected by an amount depending on the relative velocity of sound in the fluid and the flowing rate of the fluid (Fig. 21*e*). One arrangement detects this deflection with two receivers mounted on the other side of the pipe. The relative intensity of the two receivers is an indication of the flowing-fluid velocity. A change in the velocity of sound in the fluid will introduce an error.

Doppler meter. Flow rate can be detected by the Doppler shift of a reflection from particles or discontinuities in the flowing fluid. A transceiver on one side of the pipe may have an error due to the reflections from the slower particles near the wall. Figure 21*f* shows the measurement being made by opposing elements. *See* DOPPLER EFFECT.

Noise-type meter. One flow measurement device is available where the noise that is generated by the flowing fluid is measured in a selected frequency band, with a useful but relatively inaccurate readout (Fig. 21*g*).

Open channels. A different application of ultrasonics to flow measurement can be found in flumes and weirs where the level of the flowing fluid, measured by ultrasonic reflection, is used to calculate the flow rate. In unrestricted open channels the level can be measured with one reflecting transducer (either above or below the surface) and the flowing rate with a pair of submerged detectors. An electronic

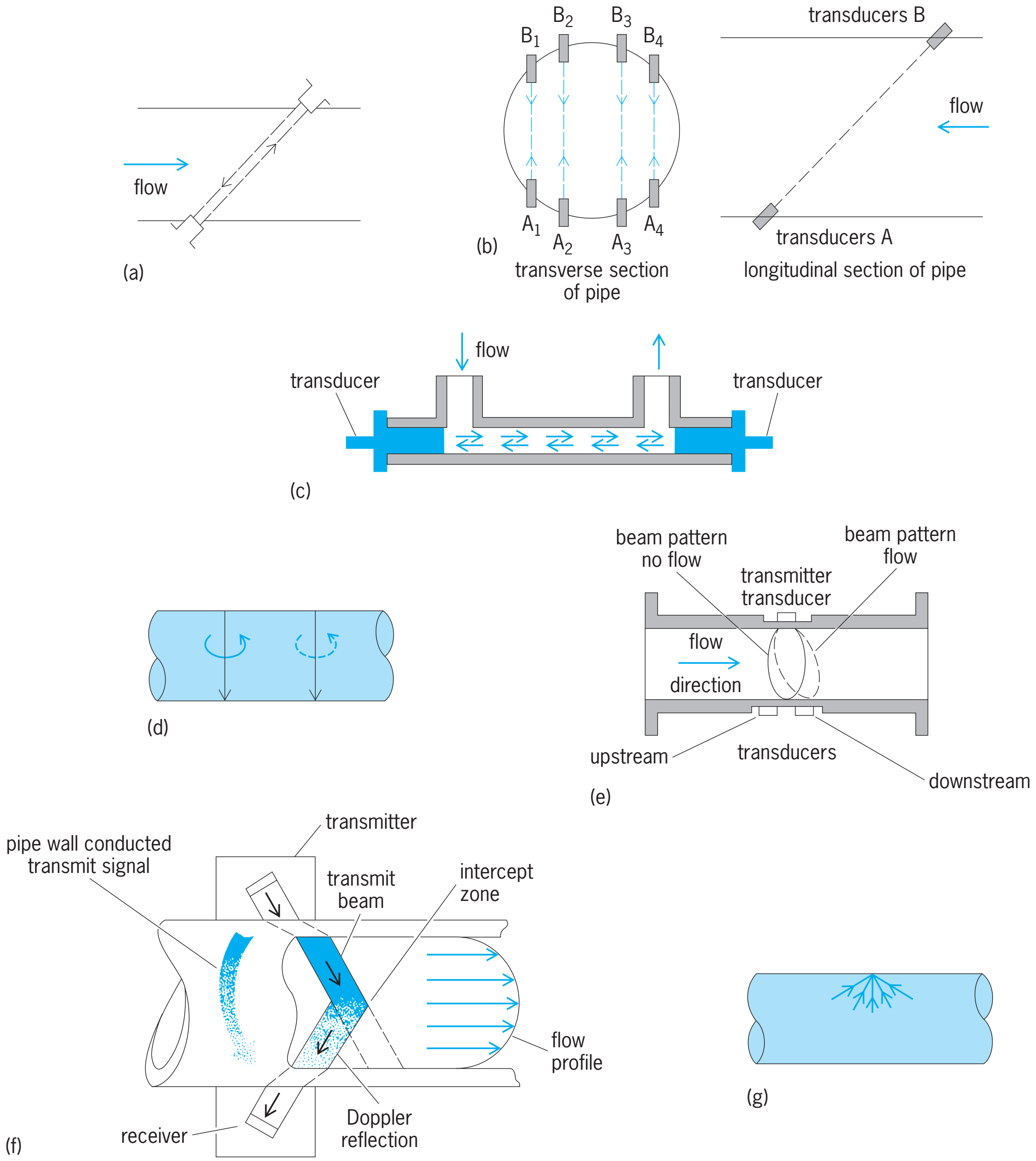


Fig. 21. Ultrasonic flowmeters. (*a*) Contrapropagating type (*after Fischer and Porter Co*., *Publ. 12257*, *1957*). (*b*) Multiple-beam contrapropagating type (*after Institute of Mechanical Engineers*, *How To Choose a Flowmeter*, *1975*). (*c*) Axial-transmission contrapropagating type (*after Dupont Co*., *Industrial Products*). (*d*) Correlation type (*after W. P. Mason and R. Thurston*, *eds*., *Physical Acoustics*, *vol. 14*, *Academic Press*). (*e*) Deflection (drift) type (*after H. E. Dalke and W. Walkowitz*, *A new ultrasonic flowmeter for industry*, *ISA J*., *7*(*10*):*60–63*, *1960*). (*f*) Doppler type (*after H. M. Morris, Ultrasonic flowmeter uses wide beam technique to measure flow*, *Control Eng*., *27*(*7*): *99–101*, *1980*). (*g*) Noise type (*after W. P. Mason and R. Thurston*, *eds*., *Physical Acoustics*, *vol. 14*, *Academic Press*)

computation incorporating the cross-section shape of the channel will give the flow rate.

Nuclear Magnetic Resonance Flowmeter

In this device (**Fig. 22**), nuclei of the flowing fluid are resonated by a radio-frequency field superimposed on an intense permanent magnetic field. A detector downstream measures the amount of decay of the resonance and thereby senses the velocity of the fluid. The most effective fluids are hydrocarbons, fluorocarbons, and water-bearing liquids because this resonance is most pronounced in fluorine and hydrogen nuclei. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Laser Doppler Velocimeter

When light is scattered from a moving object, a stationary observer will see a change in the frequency of the scattered light (Doppler shift) proportional to the velocity of the object. In the laser Doppler velocimeter (**Fig. 23**) this Doppler shift measures the velocity of particles in a fluid. From the particle velocity the fluid's velocity is inferred. A laser is used as the light source because it is easily focused and is

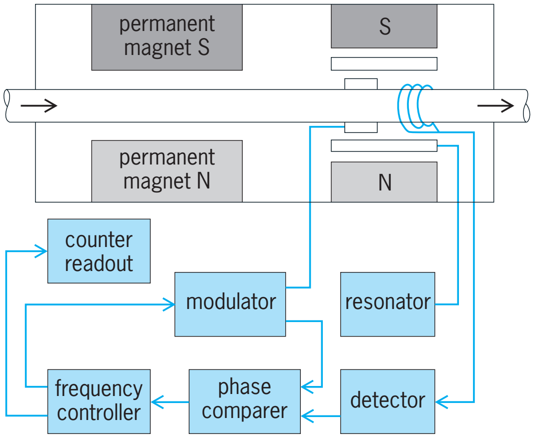


Fig. 22. Nuclear magnetic resonance type. (*After 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*)

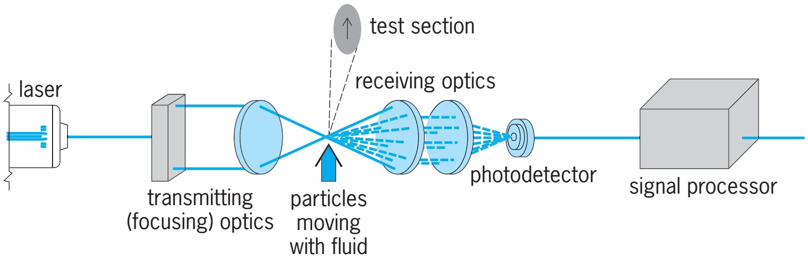


Fig. 23. Laser Doppler velocimeter. (*Thermo-Systems, Inc.*)

coherent and thereby enables the measurement of a frequency shift that is due to a Doppler effect. *See* LASER.

Accuracy of Measurement

The practical accuracy statement must be for the total flow-measurement system (combined primary and secondary devices). This has been customarily expressed as a percentage of the upper-range value ("full scale") in the case of rate meters used for industrial process control, and as a percentage of total flow rate for quantity meters. However, linear meters have often been specified in percentage of flow regardless of the use. Well-established standards for terminology, test procedures, and statements of operating environments have made it relatively easy to know what to expect for performance of secondary devices. Theory does not yet allow an accurate enough prediction of the actual flow of fluid inside of a pipe to predict the performance of a flowmeter. Square-edged orifices, venturis, and nozzles are the only types for which sufficient laboratory data have been collected and analyzed to enable a user to purchase and install a meter without individual calibration. One can expect a 1% maximum uncertainty in the primary if care is taken to follow the installation requirements of standards.

The readability or resolution of the readout display is important. This is poor at the low end of the square-root scale of differential-pressure meters. Rangeability of a flowmeter, sometimes termed turndown ratio in process applications, refers to the ratio of full-scale flow to a practical minimum flow rate. A figure of 3.5:1 is normally used for square-root scale flowmeters, while linear flowmeters, such as magnetic flowmeters, vortex-shedding flowmeters, and ultrasonic flowmeters normally have figures of 10:1 or higher. Flowmeters with rotating parts in the flowing stream, although linear, may have a smaller useful range, depending on the effect of viscosity and friction at low flow rates. Mead Bradner; Lewis P. Emerson

Bibliography. N. P. Cheremisinoff, *Flow Measurement for Engineers and Scientists*, 1987; D. M. Considine and G. K. McMillan (eds.), *Process/Industrial Instruments and Controls Handbook*, 5th ed., 1999; Daniel Industries Staff, *Fluid Flow Measurement*, 1993; J. P. DeCarlo, *Fundamentals of Flow Measurement*, 1984; C. J. Ing, *Handbook of Fluid Flowmeters*, 1988; B. G. Liptak, *Flow Measurement*, 1993; R. W. Miller, *Flow Measurement Engineering Handbook*, 3d ed., 1996; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; D. W. Spitzer, *Industrial Flow Measurement*, 2d ed., 1990.

Flow of granular solids

The gravity flow of particulate solids such as ceramic powders, plastic pellets, and ores in the form of separate particles, with the particles in contact and the voids between them filled with gas, usually air.

Particulate solids are a two-phase, solid–gas system. They are compressible; their bulk density changes during flow. Since the volume of the particles changes little during this process, it is the size of the voids that is mostly affected. Changes in the size of the voids can cause changes in gas pressure and result in gas-pressure gradients across a flowing solid which tend to reduce the rate of gravity discharge. When the solid is made up of large particles, that is, its permeability is high, when the required flow rates are low, the gas-pressure gradients are not significant; the gaseous phase can be ignored and the solid treated as a one-phase, solid-only system. The theory of flow of such a system dates to the early 1960s and has since found general acceptance. Its application to storage vessels will be outlined. This theory has been extended to include effects of the gaseous phase; this will be discussed briefly.

Bins. A bin (silo, bunker) generally consists of a vertical cylinder and a converging hopper. The size of a bin can vary from that of a small press-feed hopper with a capacity of a few quarts (liters), through portable bins, to large stationary silos storing several thousand tons of material. From the standpoint of flow, there are three types of bins: mass-flow, funnel-flow, and expanded-flow.

Mass-flow bins. Mass flow occurs when the hopper walls are sufficiently steep and smooth to cause flow of all the solid, without stagnant regions, whenever any solid is withdrawn. Valleys are not permitted, nor are ledges or protrusions into the hopper. In addition, the outlet must be fully effective, that is, if the hopper is equipped with a shut-off gate, the gate must be fully open; if it is equipped with a feeder, the feeder must draw material across the full outlet area. The range of hopper slope and friction angles leading to mass flow is shown in **Fig. 1**. Funnel flow will occur unless both the conditions for mass flow shown in this figure (the condition on θ_c in Fig. 1*b* and the condition on θ_p in Fig. 1*c*) are satisfied. The θ_c values in Fig. 1*b* also apply to conical hoppers, and the θ_p values in Fig. 1*c* also apply to wedge-shaped and chisel hoppers, provided the outlet length-to-width ratio is at least 3:1.

Mass-flow bins have advantages. Flow is uniform, and feed density is practically independent of the head of solid in the bin; this frequently permits the use of volumetric feeders for feed rate control. Indicators of low material level work reliably. Segregation is minimized because, while a solid may segregate at the point of charge into the bin, continuity of flow enforces remixing of the fractions within the hopper. Special bin design can be used for in-bin blending by circulation of the stored solid.

Mass-flow bins have a first-in-first-out flow sequence, thus ensuring uniform residence time and deaeration of the stored solid. Airlocks can often be dispensed with, provided that a critical flow rate is not exceeded; otherwise material flow will be erratic and, at high rates, material will remain fluidized and flush (pour out like a liquid) on exiting the bin.

The mass-flow type of bin is recommended for cohesive materials, for materials which degrade with time, for powders, and when segregation needs to be minimized.

Funnel-flow bins. Funnel flow occurs when the hopper walls are not sufficiently steep and smooth to force material to slide along the walls or when the outlet of a mass-flow bin is not fully effective.

In a funnel-flow bin, solid flows toward the outlet through a channel that forms within stagnant material. The diameter of that channel approximates the largest dimension of the effective outlet. When the outlet is fully effective, this dimension is the diameter of a circular outlet, the diagonal of a square outlet, or the length of an oblong outlet. As the level of solid within the channel drops, layers slough off the top of the stagnant mass and fall into the channel. This erratic behavior is detrimental with cohesive solids since the falling material packs on impact, thereby increasing the chance of material developing a stable arch across the hopper so that a complete stoppage of flow results. A channel, especially a narrow, high-velocity channel, may empty out completely, forming what is known as a rathole, and powder charged into the bin then flushes through. Powders flowing at a high rate in a funnel-flow bin may remain fluidized because of the short residence time in the bin, and flush on exiting the bin. A rotary valve is often used under these conditions to contain the material, but a uniform flow rate cannot be ensured because of the erratic flow to the valve.

Funnel-flow bins are more prone to cause arching of cohesive solids than mass-flow bins, and they therefore require larger outlets for dependable flow. These bins also cause segregation of solids (**Fig. 2**) and are unsuitable for solids which degrade with time in the stagnant regions. Cleanout of a funnel-flow bin is often uncertain because solids in the stagnant regions may pack and cake.

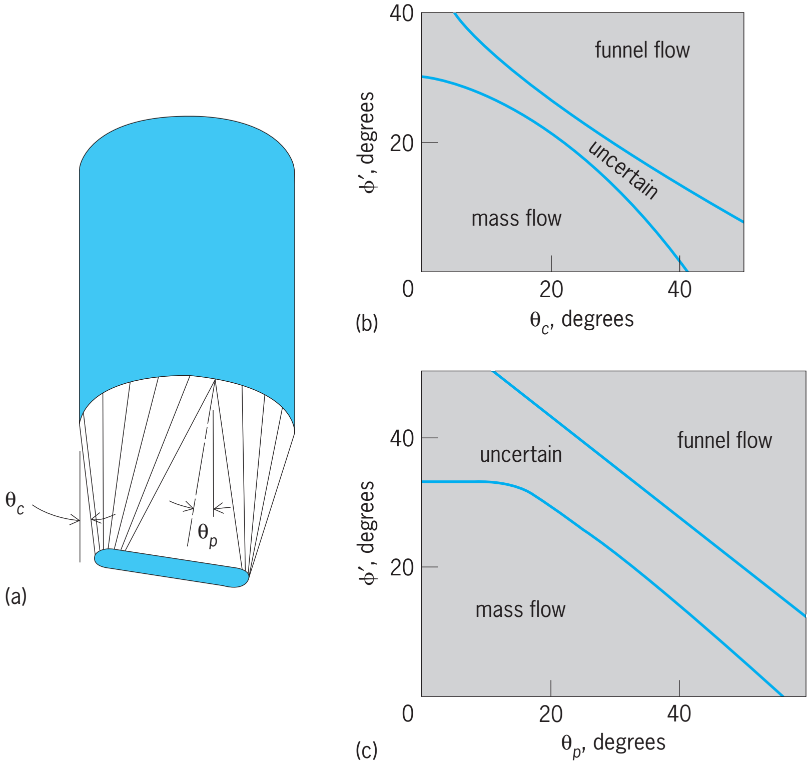


Fig. 1. Bounds on mass flow and funnel flow. (*a*) Geometry of a transition hopper, showing slope angles θ_c and θ_p. (*b*) Range of kinematic angle of friction (ϕ') and θ_c leading to mass flow and funnel flow. (*c*) Range of $\phi'$and θ_p leading to mass flow and funnel flow.

The funnel-flow type of bin is suitable for coarse, free-flowing, or slightly cohesive, nondegrading solids when segregation is unimportant.

Expanded-flow bins. These are formed by attaching a mass-flow hopper to the bottom of a funnel-flow bin. The outlet usually requires a smaller feeder than would be the case for a funnel-flow bin. The mass-flow hopper should expand the flow channel to a dimension sufficient to prevent ratholing.

These bins are recommended for the storage of large quantities [10,000 ft^3 (300 m^3) or more] of nondegrading solids. This design is also useful as a modification of existing funnel-flow bins to correct erratic flow caused by arching, ratholing, or flushing. The concept can be used with single or multiple outlets.

Feeders. The hopper outlet must be fully effective. If flow is controlled by a feeder, the feeder must be designed to draw uniformly through the entire cross section of the outlet. It is also essential that the feeder be either suspended from the bin itself or supported on a flexible frame so as to readily deflect with the bin.

Vibrators are suitable for materials which, free-flowing under conditions of continuous flow, cake when stored at rest, but break up into separate particles when vibrated—for example, granular sugar. Fine powders and wet materials tend to pack when vibrated; hence vibration is not recommended. A mass-flow bin with a screw, belt, or similar feeder is best suited to those materials.

Design for flow. The outlet dimensions of a hopper outlet must be sufficient to assure unobstructed flow. Flow may be obstructed by interlocking of large particles and by cohesive doming and ratholing across or above the outlet. Prevention of interlocking requires that the outlet be larger than several particle sizes. Prevention of cohesive obstructions requires that the outlet be sufficiently large to ensure the failure of potential cohesive obstructions.

Under the head of solid in a bin, pressure p develops within the solid. The solid consolidates and gains compressive strength f. This strength permits the development of stable domes and of stable ratholes in funnel-flow bins.

Consider a bin with a mass-flow conical hopper (**Fig. 3**). The solid is permitted to flow into it. The solid unconsolidated when deposited at the top, but, as an element of the solid flows down, it becomes consolidated under the pressure p acting on it in the bin. This pressure is shown by the p line. At first, p increases with depth, then levels off in the vertical part of the channel. At the transition to the hopper, the pressure increases abruptly and then decreases, at first exponentially and then approximately linearly toward zero at the imaginary vertex of the hopper. To each value of consolidating pressure p there corresponds a strength f of the solid, so that the strength f, generated by pressure p, similarly increases and decreases as indicated by the f line. Design for flow is based on the flow–no-flow criterion postulated as follows: a solid will flow provided the strength f which the solid develops is less than the stress s which would act in a stable obstruction to flow, $f < s$.

In a mass-flow bin, a dome across the hopper (Fig. 3) is the potential obstruction. The stress s in a dome is proportional to the span of the bin, as shown by line s. It is now observed that lines s and f intersect. Above the point of intersection, $f < s$, the flow criterion is satisfied and the solid will flow. Below that point, $f > s$, the solid has enough strength to support a dome and will not flow. The point $f = s$ determines the critical dimension B of the hopper which has to be exceeded for dependable flow.

Flowability of hoppers. As shown in Fig. 3, both the stress s in a stable dome and the consolidating pressure p are linear functions of the width of the hopper B. Since s and p are zero at the vertex of the hopper, their ratio is constant for a given hopper. This ratio is referred to as the flow factor, $ff = p/s$. It measures the flowability of a hopper. Values of ff have been computed for a wide range of hoppers.

Flowability of solids. The relation between the consolidating pressure p and the generated compressive strength f is called the flow function FF of a solid and is a measure of its flowability. The flowability of solids depends on several parameters.

Surface moisture. Solids are least free-flowing when moisture content is in the range of 70–90% of saturation. A saturated solid will usually drain in storage to

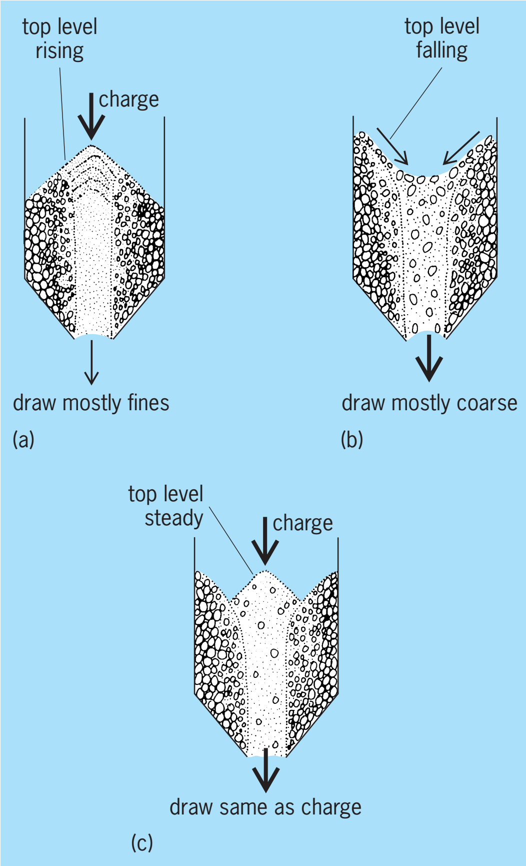


Fig. 2. Flow patterns in a funnel-flow bin, showing segregation of solids. (*a*) Charge coming in is greater than draw coming out, resulting in draw with a preponderance of fine particles. (*b*) Draw is greater than charge, resulting in draw with a preponderance of coarse particles. (*c*) Draw and charge are of same magnitude, resulting in draw with same consistency as charge. (*After A. W. Jenike, Why bins don't flow, Mech. Eng., May 1964*)

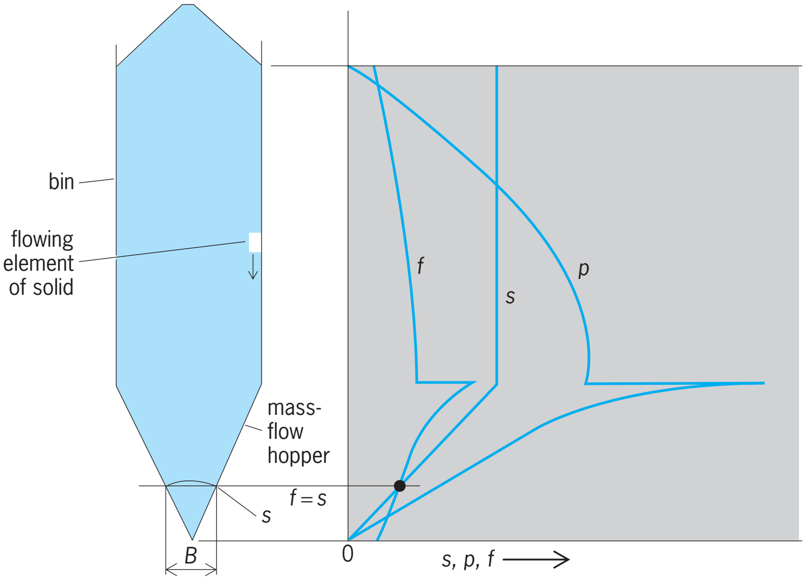


Fig. 3. Section of bin with mass-flow conical hopper, together with graph showing relative values of pressure p, compressive strength f, and stress s, and illustrating criterion ($f = s$) for minimum outlet dimension B. (*After A. W. Jenike, Quantitative design of mass-flow bins, Powder Technol., 1:237–244, 1967*)

the range of minimum flowability. Flowability tests may be carried out for a range of moisture content to determine the maximum moisture at which a solid can be handled by gravity.

Temperature. Many solids, especially those with a low softening point, are affected by temperature. The critical conditions often occur when a solid is placed in storage at an elevated temperature and allowed to cool while at rest. Some solids gain maximum strength when held at a constant elevated temperature. Flowability tests indicate the maximum temperature at which a solid can be stored and whether or not it can be cooled without excessive caking.

Time of storage at rest. Many solids are free-flowing if they are kept in motion, but cake severely if stored at rest for a period of time. Flowability tests predict the maximum time that a solid can remain in storage at rest. The most widely accepted method of measuring a solid's flowability utilizes the Jenike shear tester. Using small, representative samples (usually 5 gallons or 19 liters or less), measurements are made of the solid's cohesive strength, friction, compressibility, permeability, and other relevant properties. These tests must be run under conditions that accurately simulate how the solid is handled in an industrial setting.

Effect of gaseous phase. Incidental effects of the gaseous phase are flushing of powders and limitation of discharge flow rate. In the latter, a speedup of the bin outlet feeder produces no increase in feed rate, which is limited by the hopper outflow rate. Predicting the flow effects caused by two-phase (solid–gas) interactions is difficult. Empirical approaches are often fruitless because of the many variables involved. A combined solid–gas theoretical approach can provide the needed answers.

As briefly described above, continuum mechanics has been successfully used to describe a one-phase, solid-only system. This approach can be extended to include gaseous effects. The basic equations involve equilibrium of forces, a fluid-flow rule in addition to a solids-flow rule, and continuity of the gas and solids. By incorporating boundary conditions, it is possible to predict when flow rate limitations as well as flushing will occur. Flow rate limitations can often be overcome by enlarging the bin outlet or by introducing a small amount of gas into the bin to promote flow. In addition to applying these techniques to bin flow problems, other applications involve fixed- and moving-bed reactors, lockhoppers, and standpipes.

Inserts. Various types of devices can be placed inside the hopper section of a bin to correct flow problems. These consist of inverted cones and pyramids as well as a hopper-within-the-hopper. Some of the benefits which may accrue with such devices include converting a funnel-flow bin to mass flow, achieving completely nonsegregated discharge from a mass-flow bin, or causing the bin to act as a blender as well as a storage vessel. It is important that such devices be carefully designed and fabricated both to ensure that the desired flow pattern will develop and to maintain the structural integrity of the insert and the bin.

Segregation. Particles in a mixture can segregate by one of five primary mechanisms: sifting, particle velocity on a surface, air entrainment (fluidization), entrainment of particles in an airstream, and dynamic effects. Of these, the most common mechanism for segregation is sifting of fines through a matrix of coarse particles. For this to occur, there must be at least a minimum ratio (1.5:1) of particle sizes and a minimum mean particle size (100 micrometers), the bulk solid must be free-flowing, and there must be relative motion between particles.

In order to solve a segregation problem, it is necessary to first identify the mechanism by which the particles are segregating. Then one or more of the following techniques can be considered to reduce or eliminate segregation: changing the characteristics of the material which are causing it to segregate, changing the process sequence by which the material is being handled, and changing the design of the equipment being used to handle the material.

Abrasive wear. Abrasive wear of equipment is often an important problem when storing and handling bulk granular solids. A wear tester has been developed to measure the wear rate due to solids sliding on a wearing surface as they do during flow in a bin. A dimensionless wear ratio is measured as a function of the solids normal stress applied to the wearing surface for a particular wear material and for particular bulk solid of a certain particle size distribution, moisture content, and temperature. By incorporating the principles of bulk solids flow, it is possible to predict in absolute quantitative terms the wear which will occur in any location within a bin, hopper, feeder, chute, or other bulk-solids-handling equipment. In the design or modification of such equipment, this serves as an invaluable aid to the engineer when faced with questions of placement, quantity, type, and service life of wear-resistant materials. *See* BULK-HANDLING MACHINES; PARTICULATES; WEAR.

Andrew W. Jenike; John W. Carson

Bibliography. J. W. Carson and J. Marinelli, Characterize bulk solids to ensure smooth flow, *Chem. Eng.*, pp. 78–89, April 1994; J. Marinelli and J. W. Carson, Solve solids flow problem in bins, hoppers and feeders, *Chem. Eng. Prog.*, pp. 22–27, May 1992; J. Marinelli and J. W. Carson, Use screw feeders effectively, *Chem. Eng. Prog.*, pp. 47–51, December 1992.

Flower

A higher plant's sexual apparatus in the aggregate, including the parts that produce sex cells and closely associated attractive and protective parts (**Fig. 1**). "Flower" has been used broadly enough by some botanists to include the reproductive equipment of cycads, conifers, and other plants with naked seeds. Here, however, it will be limited, as is usual, to the angiosperms, plants with enclosed seeds and the unique reproductive process called double fertilization. In its most familiar form a flower is made up

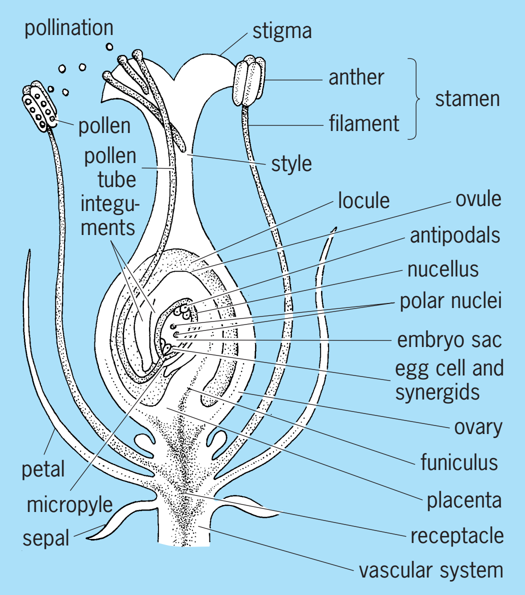


Fig. 1. Flower structure, median longitudinal section.

of four kinds of units arranged concentrically. The green sepals (collectively termed the calyx) are outermost, showy petals (the corolla) next, then the pollen-bearing units (stamens, androecium), and finally the centrally placed seedbearing units (carpels, gynoecium). This is the "complete" flower of early botanists, but it is only one of an almost overwhelming array of floral forms. One or more kinds of units may be lacking or hard to recognize depending on the species, and evolutionary modification has been so great in some groups of angiosperms that a flower cluster (inflorescence) can look like a single flower.

Floral Diversity

From the time of Carl Linnaeus, the classification of angiosperms has been based on floral features, but the diversity that makes this possible has also made it hard for botanists to develop widely applicable concepts and unequivocal terms. The old term pistil, for example, denoting the pestlelike shape of many gynoecia, could be applied to a unit or to a compound structure. It has gradually given way to carpel, which always refers to a unit. Seeking simplicity and consistency, pre-Darwinian botanists commonly reduced all the structural complexities of higher plants to two or three elements: shoot and root; axis and appendages; or root, stem, and leaf. A flower was viewed as a shortened axis bearing modified leaves. This was the "primitive" flower, the ground plan, or the archetype from which all sorts of flowers could be derived conceptually by suppressing parts (abortion), by uniting parts of the same kind (connation), or by uniting unlike parts (adnation). Interpreting the gynoecium in all its diversity was especially troublesome, and much effort went into two approaches: analyzing flower development (ontogeny) and using the position of vascular bundles to delimit the floral parts.

The terms and concepts of this era were not intended to reflect ancestor-descendant relationships. Some workers believed, rather, that regularities in plant structure would parallel regularities in crystal structure, and crystals in all their diversity had been found to be geometric modifications of a few primitive forms. Others thought the "true nature" of the flower—a notion traceable to the *physis* of pre-Socratic philosophy—could be discovered by study and contemplation. All this was brought into the evolutionary era, common ancestry replacing archetype or ground plan and "true nature" tagging along, with mixed results. Many of the floral modifications identified by pre-Darwinians were, and still are, readily seen as changes brought about by natural selection. It is now widely realized, however, that it serves no purpose to look for axes and appendages in highly altered reproductive structures and that the nature of the flower belongs to metaphysics, not science.

Though concepts have changed, the old expressions abortive, adnate, and connate survive as descriptive terms. A floral part can be called abortive if its development, when compared with that of corresponding parts in the same species or a related species, is stunted: *Ranunculus abortinus,* the abortive or kidney-leaf buttercup, has diminutive petals and carpels. Like parts of a flower are connate if they merge when followed up or down: the stamens of an orange blossom (**Fig. 2**) are basally connate, and the petals of a petunia are connate in a tubular corolla. Unlike parts are adnate if they merge: the stamens of mints, fox gloves, petunias, and many other plants are adnate to a corolla tube.

Concentric units. Most botanical terms have never been anything but descriptive, and a botanist must have a large store of them to impart the multiformity of flowers. The examples that follow are only a smattering. An extra series of appendages alternating with the sepals, as in purple loosestrife, is an epicalyx. A petal with a broad distal region and a narrow proximal region is said to have a blade and claw: the crape myrtle has such petals. The term perianth, which embraces calyx and corolla and avoids the need to distinguish between them, is especially useful for a flower like the tulip, where the perianth parts are in two series but are alike in size,

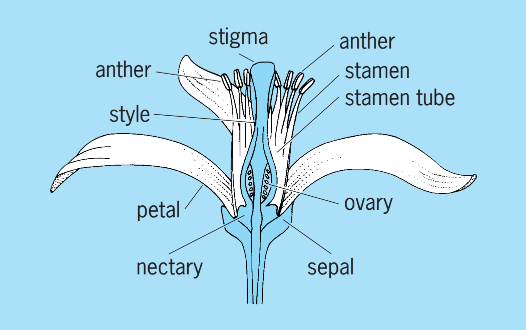


Fig. 2. Diagrammatic section of an orange blossom. (*After J. B. Free*, *Insect Pollination of Crops*, *Academic Press*, *1970*)

shape, and color. The members of such an undifferentiated perianth are tepals. When the perianth has only one series of parts, however, they are customarily called sepals even if they are petallike, as in the windflower.

A stamen commonly consists of a slender filament topped by a four-lobed anther, each lobe housing a pollen sac. In some plants one or more of the androecial parts are sterile rudiments called staminodes: a foxglove flower has four fertile stamens and a staminode (**Fig. 3**). Carpellode is the corresponding term for an imperfectly formed gynoecial unit. Most stamens release their pollen by longitudinal splitting of the anther along the anther's inner face, along its outer face, or on its sides, depending on the species (introrse, extrorse, or lateral dehiscence). In others—the stamens of rhododendrons, azaleas, or potato flowers, for instance—the pollen exits from a pair of terminal pores (poricidal dehiscence).

A gynoecium is apocarpous if the carpels are separate (magnolia, blackberry; **Fig. 4**) and syncarpous if they are connate (tulip, poppy). Or the gynoecium may regularly consist of only one carpel (bean, cherry). A solitary carpel or a syncarpous gynoecium can often be divided into three regions: a terminal, pollen-receptive stigma; a swollen basal ovary enclosing the undeveloped seeds (ovules); and a constricted, elongate style between the two. The gynoecium can be apocarpous above and syncarpous below; that is, there can be separate styles and stigmas on one ovary (wood sorrel).

Every flower cited so far has a superior ovary: perianth and androecium diverge beneath it (hypogyny). If perianth and androecium diverge from the ovary's summit, the ovary is inferior and the flower is epigynous (apple, banana, pumpkin). A flower is perigynous if the ovary is superior within a cup and the other floral parts diverge from the cup's rim (cherry). A syncarpous ovary is unilocular if it has only one seed chamber, plurilocular if septa divide it into more than one. The ovules of a plurilocular ovary are usually attached to the angles where the septa meet; this is axile placentation, a placenta being a region of ovular attachment. There are other ways in which the ovules can be attached—apically, basally, parietally, or on a free-standing central placenta—each characteristic of certain plant groups (**Fig. 5**).

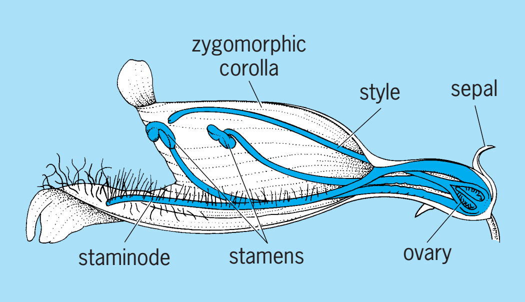


Fig. 3. Flower of *Penstemon*, a member of the foxglove family, in section. (*After C. E. Wood, Jr., A Student's Atlas of Flowering Plants: Some Dicotyledons of Eastern North America, Harper and Row, 1974*)

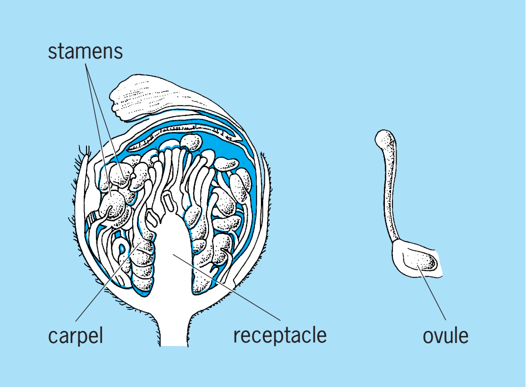


Fig. 4. Longitudinal section of a flower bud and one carpel in section of blackberry (*Rubus betulifolius*). (*After K. R. Robertson, The genera of Rosaceae in the southeastern United States, J. Arnold Arbor., 55:344–401, 1974*)

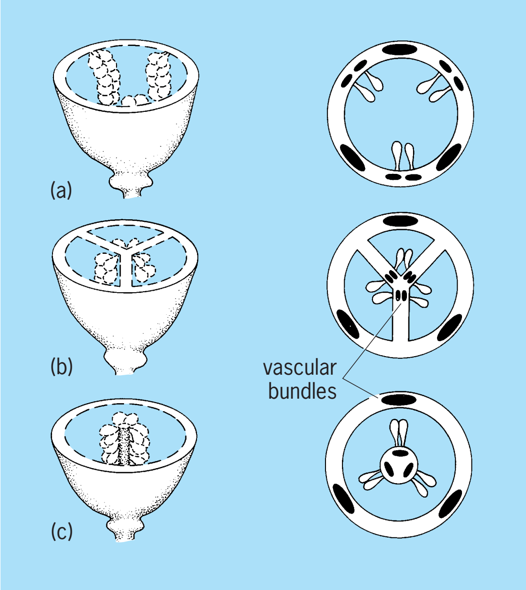


Fig. 5. Placentation: (*a*) parietal, (*b*) axile, and (*c*) free central. (*After P. H. Raven, R. F. Evert, and H. Curtis, Biology of Plants, Worth Publishers, 1976*)

The term bract can be applied to any leaflike part associated with one or more flowers but not part of a flower. Floral bracts are frequently small, even scalelike, but the flowering dogwood has four big petallike bracts below each flower cluster. The broad end of a flower stalk where the floral parts are attached is the receptacle. The same term is used, rather inconsistently, for the broad base bearing the many individual flowers (florets) that make up a composite flower like a dandelion or a sunflower.

Symmetry. A flower is actinomorphic if it is radially symmetrical, zygomorphic if bilaterally symmetrical. It is isomerous if each series has the same number of parts and it is 3-, 4-, or *n*-merous according to the number of parts in a series. Thus, an evening primrose, with four sepals, four petals, two whorls of four stamens each, and four carpels, is isomerous and 4-merous. It is also diplostemonous because its stamens are in two series; with one series it would be

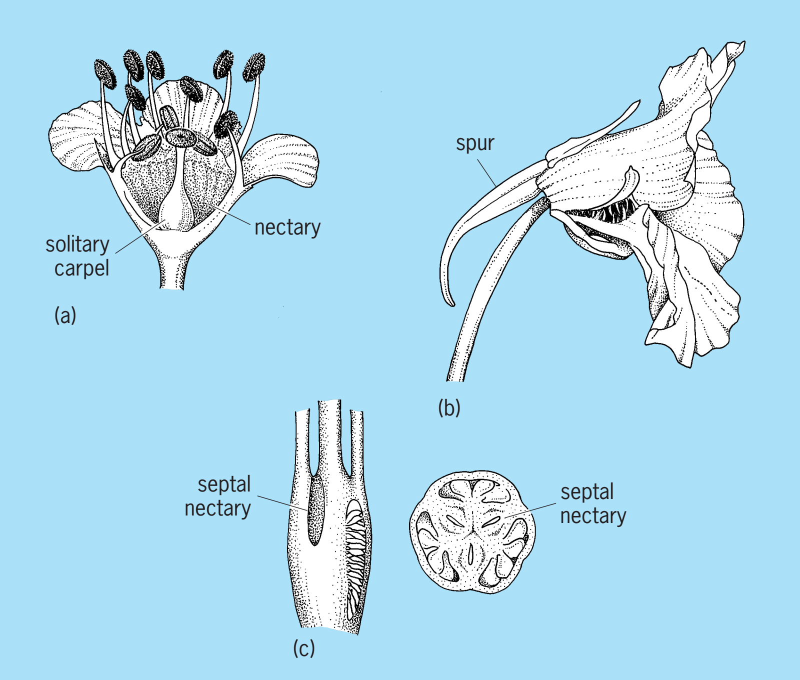


Fig. 6. Nectaries. (*a*) Cherry flower (*after W. H. Brown, The bearing of nectaries on the phylogeny of flowering plants, Proc. Amer. Phil. Soc., 79:549–595, 1938*). (*b*) Flower of nasturtium (*after B. J. D. Meeuse, The Story of Pollination, Ronald Press, 1961*). (c) Ovary of gladiolus (longitudinal and cross section) showing septal nectaries (*after W. H. Brown, 1938*).

haplostemonous. Flowers of dicotyledons are commonly 5-merous (carpels usually fewer), less often 3-merous, 4-merous, or 5+-merous; whereas monocot flowers are commonly 3-merous and rarely 4-merous (carpels often fewer), but never 5-merous.

Sexuality. A plant species is diclinous if its stamens and carpels are in separate flowers. A diclinous species is monoecious if each plant bears staminate and carpellate (pistillate) flowers, dioecious if the staminate and carpellate flowers are on different plants. The corn plant, with staminate inflorescences (tassels) on top and carpellate inflorescences (ears) along the stalk, is monoecious. Hemp is a well-known dioecious plant.

Nectaries. Flowers pollinated by insects or other animals commonly have one or more nectaries, regions of densely cytoplasmic tissue that secrete a sugar solution. A nectary can be nothing more than a layer of tissue lining part of a floral tube or cup (cherry; **Fig. 6***a*), or it can be as conspicuous as the secretory spur of a nasturtium (Fig. 6*b*) or a larkspur. It can be a cushionlike outgrowth at the base of a superior ovary (orange blossom) or atop an inferior ovary (parsley family). Gladiolus and a number of other monocotyledons have septal nectaries, deep secretory crevices where the carpels come together (Fig. 6*c*). Substances that give off floral odors—essential oils for the most part—ordinarily originate close to the nectar-producing region but are not coincident with it. Production by the epidermis of perianth parts is most common, but in some species the odor emanates from a more restricted region and may even come from a special flap or brush. Most insect-pollinated plants have visual cues, some of them outside the human spectral range, as well as odor to bring the pollinators to the flowers and guide them to the nectar. *See* SECRETORY STRUCTURES (PLANT).

Inflorescence. Inflorescence structure, the way the flowers are clustered or arranged on a flowering branch, is almost as diverse as floral structure. To appreciate this, one need only contrast the drooping inflorescences (catkins) of a birch tree with the coiled flowers of a forget-me-not or with the solitary flower of a tulip. In some cases one kind of inflorescence characterizes a whole plant family. Queen Anne's lace and other members of the parsley family (Umbelliferae) have umbrellalike inflorescences with the flower stalks radiating from almost the same point in a cluster. The stalkless flowers (florets) of the grass family are grouped into clusters called spikelets, and these in turn are variously arranged in different grasses. Flowers of the arum family (calla lily, jack-in-the-pulpit), also stalkless, are crowded on a thick, fleshy, elongate axis. In the composite family, florets are joined in a tight head at the end of the axis; the heads of some composites contain two kinds (**Fig. 7**), centrally placed florets with small tubular corollas and peripheral ray florets with showy, strap-shaped corollas (the "petals" one plucks from a daisy). Inflorescences differ so much with regard to patterns of branching, the order in which flowers appear, and the distribution of bracts, that a system for describing and comparing all of them is almost beyond reach. There have been a few attempts at

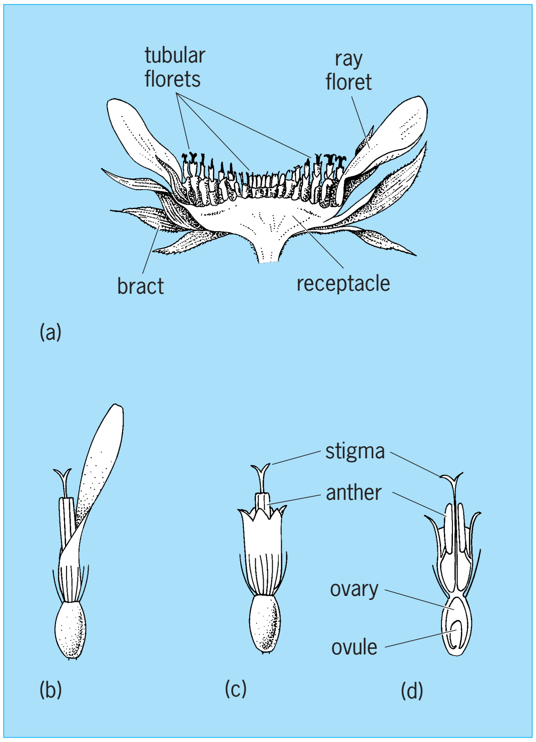


Fig. 7. Sunflower (*Helianthus annuus*) of the composite family. (*a*) Inflorescence in section. (*b*) Side view of ray floret. (*c*) Side view of tubular floret. (*d*) Section of tubular floret. (*After J. B. Free, Insect Pollination of Crops, Academic Press, 1970*)

computer modeling, with some success. *See* INFLORESCENCE.

Evolution

Inferring how flowers began is another of botany's troublesome tasks. Ancient impressions of flowers are rare and usually short on structural detail. Though flowers embedded in amber can be exquisitely preserved, none is old enough to illuminate the origin of angiosperms. Remains of dispersed pollen grains, on the other hand, are plentiful enough and old enough, and with the help of the electron microscope angiosperm pollen can be distinguished from gymnosperm pollen well enough to track the entry of angiosperms into the record. Investigations of such fossils place the beginning of flowering plants in the Early Cretaceous and support the orthodox view that they came from a common ancestral group.

Traits of early flowers. Though the ancestors of angiosperms still cannot be identified among fossil gymnosperms, the record supports the opinion that early flowers had traits seen today, in various combinations, in such living members of the order Magnoliales as the star anise (*Illicium*), the Fijian tree *Degeneria,* and *Drimys,* a genus of the Old and New World tropics. The early flowers were bisexual, with floral parts indefinite in number, free from each other, and probably arranged in low spirals. The perianth was not differentiated into calyx and corolla. The carpel may have been folded in half lengthwise along its midvein, with a stigmatic crest running along its unsealed margins. *Drimys piperita,* a magnolialean shrub of the western Pacific, has such carpels, and they are duplicated to some extent in a fossil from the Cretaceous of central Kansas.

It has been recognized since the early 1800s that all flowers can be derived conceptually from a magnolialean archetype. Twentieth-century anatomists gave the concept an evolutionary basis by pointing out, among other things, that the stamens and carpels of some geographically remote Magnoliales can be arranged as plausible vestiges of early evolutionary change. Pollen studies support the anatomical findings, demonstrating that pollen grains with single furrows, a trait found in archaic Magnoliales, are the oldest in the angiosperm record, and other forms enter the record in younger horizons as clear-cut indications of evolutionary diversification. *See* PALYNOLOGY.

Diversification and coevolution. The fossil record shows that diversification of angiosperms was rapid. It took perhaps 25,000,000–30,000,000 years for flowering plants to appear, diversify, and gain dominance over the gymnosperms, about as much time as it later took for mammals to radiate and replace reptiles and the dominant animals. It is thought that the earliest mammals, along with birds and insects, brought the flowering plants to dominance. The first angiosperms were plants of limited distribution, low abundance, and low diversity, with seeds dispersed rather inefficiently by wind, water, and reptiles. The emergence of fruit-eating and seed-eating birds and mammals enabled undigested seeds to get farther from the parent plant and its predators before germinating, thereby spreading out the species. This in turn gave insect pollination an advantage over wind pollination, which is the usual method for gymnosperms, because individuals of the same species were now often too far apart for wind-mediated outbreeding. Thus, flowering plants began to coevolve with the insects that had heretofore visited haphazardly to feed on pollen or on the flower.

Pollination. Beetles, an old group, are probably the original pollinating animals, and some of the Magnoliales have floral adaptations that perpetuate this form of breeding. In many other plants, however, pollination by beetles has given way to pollination by bees, wasps, butterflies, moths, birds, bats, or other animals. Regardless of the pollinator, the flower ordinarily offers nectar as a reward, but the way in which the reward is advertised and protected differs, depending on the plant's major pollinator. Thus, bee flowers can be bright orange, yellow, purple, or blue; but they are rarely red, a color that bees cannot see. Bird flowers, on the other hand, are often bright red, eliminating bees as competitors for the nectar. A bee flower's corolla may form a gullet above the nectary in such a way that a bee can enter but its competitors cannot; the stamens may be placed so as to release pollen onto the bee when it squeezes into the gullet; and the gullet may be so formed as to foil honeybees while admitting bumblebees (foxglove). Bee flowers that offer only pollen are constructed differently from those that also offer nectar.

Butterfly flowers and moth flowers are fragrant like most bee flowers, but they commonly have long, narrow, bee-proof tubes above the nectary that are well suited to the long, coiled lepidopterous tongue. Flowers pollinated by butterflies are brightly colored, open in daylight, and often end in a broad landing platform (petunia). Flowers pollinated by nocturnal moths are pale, and in keeping with the hovering habit of such insects they usually lack a landing platform.

Bird flowers commonly have tubular corollas that are larger in diameter than the corollas of moth flowers and butterfly flowers. They are rarely scented, since birds have little or no sense of smell. The ovary is often inferior; the nectary above it can then be probed by a bird's bill without damage to the ovules. A typical bat-pollinated flower is a large, pallid night-bloomer with many anthers and a strong odor, frequently described as musty or rancid. Bird flowers and bat flowers are all heavy nectar producers, for the high energy needs of the animals must be met to sustain the relationship.

Some flowers have coevolved with flies that eat pollen and nectar. Some of them use shiny surfaces or dangling parts to bring in the pollinators. Other flowers (or flowerlike inflorescences) attract eaters of carrion or dung by producing an odor resembling the pollinator's favorite food; they may be constructed so as to trap flies or carrion beetles for a time, thereby increasing the chance that the insect will take on some pollen before departing.

This brief overview hardly begins to convey the array of modifications that evolution has wrought in the form and physiology of flowers. In a number of lin eages the flowers have clustered so that the inflorescence functions as one flower in attracting or powdering the pollinator. Dogwoods and the huge family Compositae are examples. In some tropical tree species, flower clusters do not appear at the ends of branches but burst through the bark of the branches or trunk, an adaptation to pollination or dispersal by bats. Insect pollination has been replaced by wind pollination in some groups, including the grasses, the sedges, certain composites (ragweed), and such catkin-bearing trees as the birches and walnuts. Wind-pollinated flowers often lack color, smell, and perianth parts. They are heavy producers of pollen, and the grains are usually smooth and dry in contrast to the sticky, sculptured grains of most insect-pollinated plants, though wind-pollinated stigmas are likely to be sticky or feathery. Self-pollination, with or without the help of a pollinator, is widespread among angiosperms. There are species of violets with two kinds of flowers on each plant: the insect-pollinated ones, picked for nosegays, and smaller self-pollinated ones that never open. Though some plants have mechanisms to keep seeds from developing in a self-pollinated flower, others are constructed so that successful selfing will take place if outcrossing does not. Indeed, alternative routes to reproductive success are commonplace, and the link between a flower and a pollen vector is seldom so tight that it cannot alter with a change in selection pressure. *See* POLLINATION.

Outcrossing. Outcrossing is basic to the normal evolutionary process, so it is understandable that many lineages have evolved ways to favor or enforce it. Stamens and carpels can be kept apart in unisexual flowers (dicliny) by structural barriers within a flower or by maturing at different times (dichogamy); the barrier to selfing may also be physiological, as when pollen will not germinate except on the stigma of a genetically different individual within the species.

Formation

A plant is ready to flower when it has met complex requirements that vary from one species to another, even from one geographic race to another. There must be a prior period of growth, which is short for herbs of the desert or of alpine meadows but may be decades for a tree or a century or more for a bamboo. Some long-lived plants grow for years, flower once, and die after the fruits ripen, much as an annual plant does. Seasonal requirements are known to all gardeners; a plant may need a period of exposure to low temperature, a particular day-length, or both before it will bloom. Cabbages and some other biennial plants will not bolt—send up a flowering shoot—without first experiencing a cool period. In monsoon regions, the flowering of certain trees is triggered by the onset of the dry season. Tulips and daffodils do not need a cold period for flower formation, but the initiation of flowers in the summer and their emergence from the bulb the following spring are linked to seasonal warming. Depending on the species, plants of the wet tropics can flower continuously or in response to an internal "clock" or in response to unusual weather. Though flowering intervals vary greatly in a tropical rainforest, widespread individuals of a species frequently flower at the same time; this enables the plants to cross-pollinate and ensures that seeds will be plentiful enough for some to escape their predators.

To initiate flowers instead of leaves, a growing tip changes from an indeterminate meristem to a determinate meristem, an aggregation of dividing cells that will go through a stepwise program and then stop dividing. One of the first reactions to the flowering stimulus is an increase in ribonucleic acid (RNA) synthesis. The number of ribosomes and the rate of cell division rise, along with the relative amount of dispersed nuclear chromatin, an indication that genetic information is being transcribed. Eventually, deoxyribonucleic acid (DNA) also increases. With still more cell divisions the apex changes shape, usually becoming broader and shallower than the vegetative apex. Cells deep beneath the surface acquire bigger vacuoles, and the apex then consists of a meristematic mantle over a vacuolated core. At this stage an apex programmed to become an inflorescence looks much the same as one that will be a flower. Each floral part begins with periclinal divisions (new walls parallel to the surface), usually in the second or third cell layer. As anticlinal divisions are added, a bump forms on the apex, or several bumps for all the members of a set. The usual formative sequence is acropetal—sepals, petals, stamens, carpels—but in purple loosestrife (to cite just one exception) the inner stamens and the petals start after the gynoecium is initiated. Stamen primordia become four-lobed shortly after their initial protrusion. Carpels typically begin as protuberances, then become horseshoe-shaped before taking on the form they have when they are ready for pollination. Some syncarpous gynoecia, however, begin as circulary or elliptic ridges in which primordial carpels cannot be delimited (**Fig. 8**). Continuity of parts—connation and adnation—can come about in several ways, including interprimordial growth, extension and fusion of meristems, and initiation of a new meristem in the common base below two or more primordia. Or the surfaces of adjoining parts can unite as they grow: the tips of the paired carpels of a garden periwinkle fuse by modifying the walls of juxtaposed epidermal cells. A similar surface fusion of the margins lower down in the carpel occurs in many plants. *See* APICAL MERISTEM.

Surgical experiments show that the cells of a floral apex become increasingly committed to their final form as development progresses. Bisected early enough, an apex forms two flowers. That is, sepals and other parts form around each half. When the apex is bisected after a critical stage, however, sepals form only in the regions already committed to sepal formation, and there is another critical stage for each of the other sets of floral parts. Hormones are thought to play a major role in sequential differentiation because floral development can be changed greatly by experimental application of hormones. Hormones can also alter sex expression in diclinous plants: in the cucumber, potentially staminate flowers can be made to develop as carpellate flowers. *See* PLANT HORMONES; VERNALIZATION.

Anatomy

Some of the general anatomical features of leaves can be found in the floral appendages. A cuticle-covered epidermis overlies a core of parenchyma cells in which there are branching vascular bundles (solitary bundles in most stamens). Sepal parenchyma and petal parenchyma are often spongy, but palisade parenchyma occurs only rarely in flowers and then only in sepals. As in other parts of the plant, color comes mostly from plastids in the cytoplasm and from flavonoids in the cell sap. Cells of the petal epidermis may have folded side walls that interlock so as to strengthen the tissue. In some species the outer walls of the epidermis are raised as papillae; apparently, this is part of the means of attracting pollinators, for the papillae are light reflectors. *See* LEAF.

Stamen. As a stamen develops, periclinal divisions in the second cell layer of each of its four lobes start a sequence that will end with the shedding of pollen (**Fig. 9**). The first division makes two cell layers. The outer daughter cells give rise to the wall of the pollen sac, and the inner ones (primary sporogenous cells) are destined to become pollen after further divisions. When mature, a pollen sac typically has a prominent cell layer, the endothecium, just below a less distinctive epidermis. The inner wall and the side walls of an endothecium cell carry marked thickenings, but the outer wall does not. Splitting of the ripe anther is due partly to the way in which these differently thickened walls react to drying and shrinking and partly to the smaller size of the cells along the line of splitting.

Pollen development. Primary sporogenous cells become pollen mother cells by enlargement or by division followed by enlargement. Through two meiotic divisions each mother cell becomes a tetrad of microspores (uninucleate pollen) with the haploid number of chromosomes. Cell walls disintegrate in the process, and a covering of callose, a polysaccharide, takes their place around the now rounded cells. Ordinarily, tetrads separate into monads before the anther opens, but some plants shed paired grains (*Scheuchzeria*); others shed tetrads (cattails, rhododendrons); still other shed polyads (certain acacias); and pollinators of many orchids pick up the agglutinated contents of a whole anther. While the sporogenous cells are changing to microspores, the pollen sac becomes lined with a densely cytoplasmic, sometimes multinucleate tissue, the tapetum, that nourishes the future pollen grains as they differentiate. Tapetal cells lose their walls in some groups of plants, and the protoplasts then move amebalike among the developing grains. A microspore becomes a pollen grain when its nucleus divides, yielding a generative nucleus and a tube nucleus. Membranes form

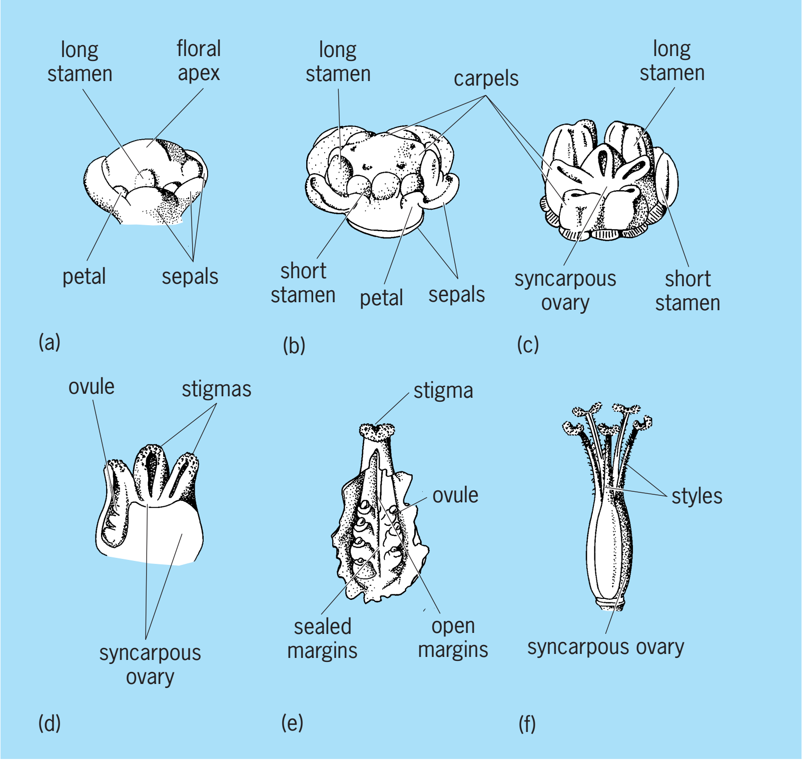


Fig. 8. Successive stages of flower formation in a wood sorrel. (*a*) Sepals and petals already initiated; the androecium begins as five bumps on the floral apex; five more will be added. At maturity, five long stamens will alternate with five short ones. (*b*) Carpel primordia taking horseshoe shape. Petal primordia lag while stamens grow. (*c*) Stamens lobed, outer parts cut away. (*d*) Gynoecium cut longitudinally. (*e*) One carpel opened from outside. (*f*) Mature gynoecium. (*After J. B. Payer, Trait d' Organogénie Comparée de la Fleur, Wheldon and Wesley, 1966*)

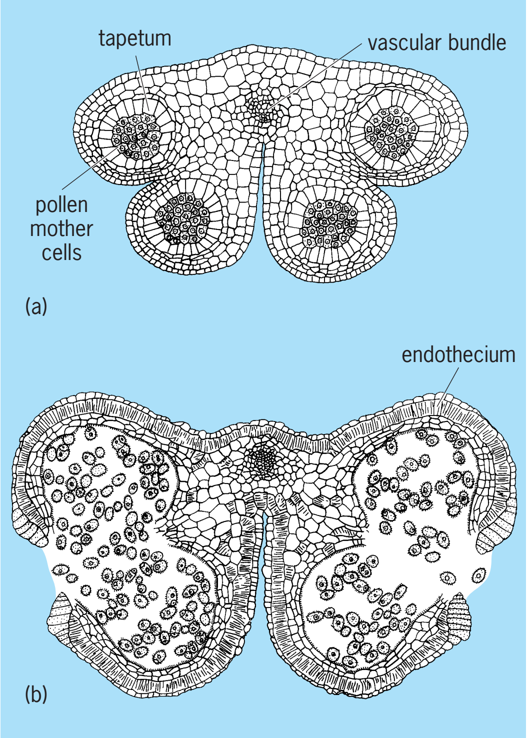


Fig. 9. Lily anther. (*a*) Cross section of young anther showing four microsporangia with sporogenous tissue. (*b*) Cross section of mature anther with two pollen sacs containing pollen. (*After A. W. Haupt, Plant Morphology, McGraw-Hill, 1953*)

around these nuclei (and in some cases a wall separates them), making the pollen grain two-celled. There are groups in which the generative nucleus cleaves almost immediately into two sperm nuclei, and the pollen is then three-nucleate, but it is more common for this final division to occur in the elongating pollen tube after the pollen germinates.

Pollen features. The pattern for outer features of the pollen grain, which vary greatly from one kind of plant to another, is set while the microspores are still encased in callose. Endoplasmic reticulum, differentially distributed under the outer membrane of each cell, determines where the germination apertures will be: no outer wall (exine) forms at the aperture sites. A temporary outer wall traversed by rods is laid down between these sites, and the addition of more material may then link the rods together. Eventually the microspores enlarge, the callose disappears, and the outer wall is permeated with sporopollenin, an oxidative polymer of carotenoids and carotenoid esters. Sporopollenin's resistance to decay and the pollen wall's range of taxonomically useful features make pollen an invaluable guide to evolution and to past environmental conditions. *See* POLLEN.

Carpel. Like other floral parts, a carpel is made up of epidermis, parenchyma, and vascular tissue. In addition, a carpel commonly has a special tissue system on which pollen germinates and through which, or along which, pollen tubes are transmitted to the ovules. Most angiosperms have solid styles, and the transmitting tissue is a column of elongate cells whose softened walls are the medium for tubal growth. The epidermis at the stigmatic end of a carpel usually changes to a dense covering of papillae or hairs; the hairs can be unicellular or pluricellular, branched or unbranched. In taxa with hollow styles, the transmitting tissue is a modified epidermis running down the stylar canal. There are two kinds of receptive surfaces, and they are distributed among the monocotyledons and the dicotyledons with taxonomic regularity. One kind has a fluid medium for germinating the pollen, and the other has a dry proteinaceous layer over the cuticle. The proteins of the dry stigmas have a role in the incompatibility reactions that encourage outbreeding.

A solitary carpel commonly has three main vascular bundles: a midvein and two marginal veins. It was realized more than a century ago that the bundles can often be used to delineate the carpels of compound gynoecia if other indicators, such as styles and placentas, are ambiguous. This works because vascular strands begin to form when the gynoecial primordia are quite small, and they can continue to differentiate as bundles while other primordial features are obscured by subsequent developmental changes. That is to say, a carpel bundle can be something of an ontogenetic marker. Some botanists built on this, as organic evolution became widely accepted, to claim that vascular bundles—all vascular bundles—are phyletically conservative; that they tend to lag in various ways as the outer form undergoes evolutionary changes; that a vestigial midvein, for instance, can mark the position of a lost appendage. And some individuals, who were unfamiliar with evolutionary theory or uncomfortable with it, or more comfortable using words from old books, saw bundles as a clue to a flower's "nature."

The view that bundles are conservative underlies most American research on floral anatomy during the 1930s, 1940s, and 1950s. It has since been questioned as a generalization. That is not to say there are no lagging or vestigial bundles: an evolving lineage might very well keep an old floral bundle in place if the bundle continues to do something useful for the plant.

Ovule. Ovule development usually takes place as the gynoecium forms, but it may be retarded when there is a long interval between pollination and fertilization (oaks, orchids). A typical ovule has a stalk (funiculus), a central bulbous body (nucellus), and one or two integuments (precursors of seed coats), which cover the nucellus except for a terminal pore (micropyle). Orientation of the ovule varies from group to group. It can be erect on its stalk or bent one way or another to differing degrees. There are also taxonomic differences in the extent to which the ovule is vascularized by branches from the gynoecial vascular system.

Fertilization

The union of an angiosperm egg and sperm takes place in an embryo sac that develops within the ovule in one of several ways.

Embryo sac. The most common way is for a large cell near the micropylar end of the nucellus to divide meiotically to form a lineup of four haploid megaspores. Three of the megaspores deteriorate; the one farthest from the micropyle enlarges and divides; and the two daughter nuclei part, one moving to the micropylar end of the cell, the other to the opposite end. Two subsequent divisions and further enlargement (at the expense of the nucellus) make an eight-nucleate embryo sac with four nuclei at each pole. Three of the nuclei at the micropylar end then differentiate further and acquire membranes to form the egg apparatus, a triplet consisting of a central egg cell with two flanking cells called synergids. Three of the four nuclei at the other end become antipodal cells. The remaining nucleus (polar nucleus) from each foursome moves to the middle of the sac, pairing with its counterpart, but cell membranes do not form. The mature sac consequently has eight nuclei but only seven cells (**Fig. 10***a*). The two polar nuclei may fuse to become a diploid secondary endosperm nucleus before fertilization. The simplified sequences just described produces the *Polygonum* type of embryo sac. There are others that differ with regard to such features as the number of megaspores and the number and distribution of the nuclei in the mature sac. These differences, as well as differences in the thickness of the nucellus and the number of integuments, can be important taxonomic characters. Seen from the standpoint of alternation of

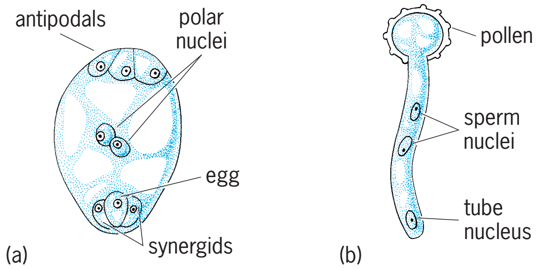


Fig. 10. Gametophyte: (*a*) embryo sac and (*b*) pollen tube. (*After O. Tippo and W. L. Stern*, *Humanistic Botany*, *W. W. Norton*, *1977*)

generations, one of the unifying themes of botany, an embryo sac is the megagametophyte of the angiosperms; a pollen tube with its contents is the microgametophyte.

Pollen tube. Emergence of the tube from the pollen grain (Fig. 10*b*) occurs as soon as the grain alights on the stigma (sugarcane), after a few minutes (dandelion), or after a longer interval—as long as two days in *Garryaf*. The grain expands by absorbing liquid from the stigma, and the plastic inner layer (intine) of the pollen wall begins to bulge through one of the apertures, where there is no exine above it. If more than one tube emerges, a characteristic of some groups, only one makes progress.

Following the transmitting tissue, tubes grow at the rate of a few millimeters per hour, usually reaching the ovule within 48 h or less, but the delay between pollination and fertilization can be weeks, even months, in some plants. The tube passes through the micropyle (in most cases), penetrates the embryo sac, and releases two sperm cells from its tip. One sperm unites with the egg cell to form a zygote, which after additional divisions becomes an embryo. The other sperm fuses with the polar nuclei or with their fusion product, the secondary endosperm nucleus; subsequent division of the resulting triple fusion nucleus makes endosperm tissue (triploid if the embryo sac is of the *Polygonum* type), which nourishes the developing embryo. Ordinarily, pollination triggers the first cell divisions toward the changing of a gynoecium to a fruit, but fertilization and seed development are required for a fruit to achieve its full growth. *See* REPRODUCTION (PLANT).

Utilization

Although human life depends on the consumption of fruits and seeds, eating the flower that precedes a fruit is something of a rarity. Pickled capers and dried cloves are flower buds, and saffron is the dried stigmas of a crocus. Honey is floral nectar, concentrated and altered by its sojourn in a beehive. Flowers are used in quantity by the perfume industry, even though many fragrances can now be synthesized. The essential oils are removed by pressing, by distillation, by extraction with a solvent, or, as in the case of highly valued jasmine, they may still be extracted by enfleurage, laying the flowers in cold tallow or lard. Raising and exporting cut flowers is a major industry in the Netherlands, in some Latin American countries, and in California and Florida; several European countries and many areas in the United States have research centers for floriculture. Thus, the flower as ornament clearly has a place in today's technological world, but to prize a flower for its beauty is not new. Archeologists have found remnants of bouquets, including bachelor's buttons and grape hyacinths, in a Neanderthal grave more than 50,000 years old. *See* FRUIT; MAGNOLIOPHYTA.

Richard H. Eyde

Bibliography. K. Esau, *Anatomy of Seed Plants*, 2d ed., 1977; A. Hallam (ed.), *Patterns of Evolution as Illustrated by the Fossil Record*, 1977; A. H. Halvey (ed.), *CRC Handbook of Flowering*, 5 vols., 1985; P. H. Raven, R. F. Evert, and H. Curtis, *Biology of Plants*, 6th ed., 1998; A. J. Richards, *Plant Breeding Systems in Seed Plants*, 1986; D. L. Strider (ed.), *Diseases of Floral Crops*, 2 vols., 1985.

Fluid coating

The operation of depositing a liquid film on a solid, due to a relative motion between them (**Fig. 1**). It occurs in many situations of everyday life as well as in numerous industrial processes. Examples include a water drop sliding on a window and leaving behind a trail; an object or a person coming out wet from a bath; or the emptying of a solid full of liquid (such as a glass containing water or a porous medium filled with oil), leaving some liquid inside the solid. In all these cases, an important practical question is the thickness of the film adhering to the solid. It may be desirable to control the thickness for coating purposes, or conversely to avoid a film altogether to keep the solid dry. To achieve this control, it is important to understand the parameters affecting the existence of such a film, and the laws of deposition.

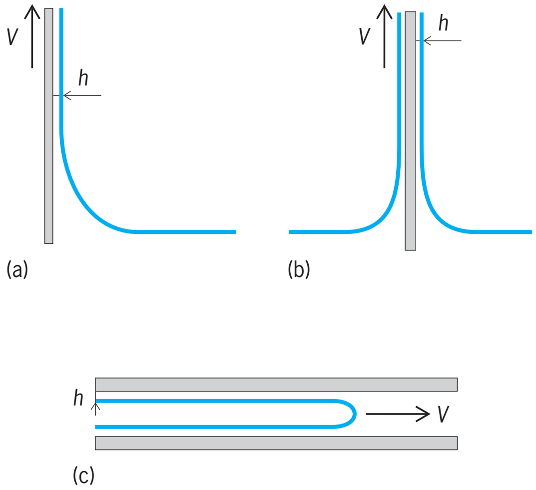


Fig. 1. Different kinds of fluid coating, according to the geometry. (*a*) Plate coating. (*b*) Fiber coating. (*c*) Film deposition in a tube. In each case, the macroscopic film coating the solid comes from a relative motion between the liquid and the solid, and the film thickness *h* can be tuned by the coating velocity *V*.

Before addressing a dynamic problem, it is often worth describing the statics. The very general situation in which a solid, a liquid, and a vapor meet together occurs, for example, in the case of a drop deposited on a solid with air all around. The different surface energies involved in such a situation (between the solid and the liquid, the liquid and the vapor, and the solid and the vapor) impose an equilibrium condition at the contact line where the three phases coexist. Since each surface tension tries to draw the contact line so as to reduce the corresponding interface, the balance of the tensions, which defines the equilibrium, fixes a contact angle close to this line. For a drop, it is the angle with which the liquid-vapor interface meets the solid. A liquid totally wets a solid if it joins it with a zero contact angle. If that is not the case (contact angle larger than zero), the corresponding situation is referred to as partial wetting. *See* SURFACE TENSION.

Total wetting can occur by itself (for example, oils on most solids) or be realized by adding wetting agents to the liquid (the most common ones being surfactants, for example, soaps in water). Drawing a solid out of such a liquid always produces a film deposition. Of course, this deposition is favored by the wetting situation, since a film would spontaneously spread from the bath onto the solid, even at rest. But such wetting films are microscopic, which is not generally the case in fluid coating (at least, for sufficiently large coating velocities): the trail of a drop is visible. The thickness of the coating film can be determined with the help of a canonical law of hydrodynamics: Because of its viscosity, a liquid in the close vicinity of a solid moves at the same velocity as the solid. This principle is the basis of fluid coating and the cause of entrainment, and calculating the film thickness requires identifying the extent of this vicinity. Different answers exist, depending on the velocity of deposition. *See* BOUNDARY-LAYER FLOW; SURFACTANT; VISCOSITY.

Low-velocity coatings. The case of low velocities was understood in the 1940s by the Russian physicists Lev Landau, Benjamin Levich, and Boris Derjaguin. The Landau-Levich-Derjaguin (LLD) picture is as follows: Drawing a solid out of a bath implies a deformation of the free surface (Fig. 1*a*), which the surface tension of the liquid opposes. The film thickness thus results from a balance between the effects of viscosity (which causes a macroscopic entrainment of liquid by the solid) and surface tension of the liquid, which resists to this deformation (and thus to the film entrainment). A dimensionless number called the capillary number compares the intensity of these antagonistic effects (defined as $\mathrm{Ca} = \eta V/\gamma$, with η and γ the liquid viscosity and surface tension, and V the deposition velocity). In the limit of low velocities, the film thickness, h, should be a function of this unique parameter (this function was found by Landau, Levich, and Derjaguin to obey a simple power law: $h \sim \mathrm{Ca}^{2/3}$), whatever the geometry of the coating. This is indeed the case, and the geometry enters only through the length that normalizes the film thickness: For plate coating, this length is the capillary length (a millimetric length that corresponds physically to the size of the static meniscus, formed when a wetting liquid meets a vertical plate), or the capillary radius for a wire or a capillary tube.

Coating at higher velocities. The hydrodynamics of fluid coating becomes more complicated at higher velocities. Then, inertia must be taken into account—in the above example, the inertia of the fluid set in motion by the solid in the bath. When the moving fluid comes out from the bath, in the tiny zone where the film thickness fixes, its inertia must be compared with the capillary force (which was found to limit the film thickness). A negligible inertia defines the LLD frame. If the inertia is not negligible, it should act to make the film thicker than predicted by the LLD picture, since it tends to pull the liquid out of the bath. It was shown experimentally in fiber (or wire) coating that this is indeed the case. The film thickness displays critical behavior as a function of velocity: Above a threshold in velocity (typically of order 1 m/s or 3 ft/s), a simple doubling of the velocity causes the film thickness to increase about tenfold (**Fig. 2**).

At still larger velocities, the film thickness stops increasing with the coating speed and even slowly decreases with it (Fig. 2). This regime concerns most industrial coating processes, where the coating speed can be of order 10–30 m/s (30–100 ft/s). At such high speeds, the contact time of a piece of solid in the bath becomes very short (it is simply $t = L/V$, with L the bath length). Thus, the problem now is to understand how much liquid this piece of solid can put in motion inside the bath—a classical question, called the viscous boundary layer problem and solved in the early twentieth century by Ludwig Prandtl. Balancing inertia with the viscous force gives the answer: The thickness of the layer moved by the solid (and thus coming out with it) increases as the square root of the contact time (it scales as $\sqrt{\eta t/\rho}$, with ρ the liquid density). Hence, the film thickness

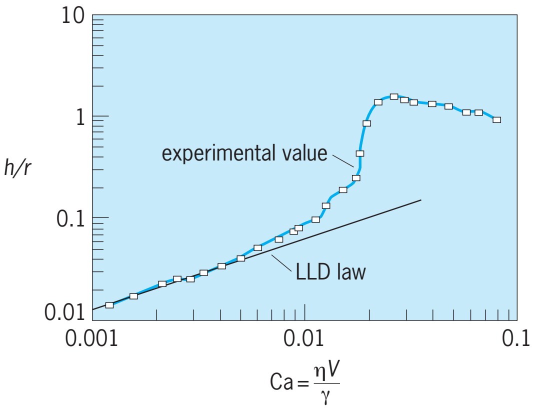


Fig. 2. **Different regimes of fluid coating, according to the coating speed. These experimental data correspond to fiber coating. The film thickness *h* normalized by the fiber radius *r* is plotted as a function of the capillary number Ca. Experimental values are compared with the Landau-Levich-Derjaguin (LLD) law for low-velocity flow.**

in this regime indeed slowly decreases with the coating speed (as $1/\sqrt{V}$, since t varies as $1/V$). This law expresses a balance between the viscosity (which is, of course, still responsible for the coating) and inertia—but now, the inertia of the bath, which is globally at rest and thus resists the motion. The action of the liquid tension, which was predominant at low velocity (the LLD frame), has become negligible in this high-speed regime.

Complex fluid coatings. Most liquids used in coating processes are complex fluids: emulsions (which are fine dispersions of oil in water) or water-containing surfactants. In the cases where inertial effects are dominant (high-speed regimes), the above description remains valid because the parameters fixing the coating are bulk parameters (viscosity and density). The situation is much more interesting at smaller velocities, where surface effects are important. For example, a soap bubble is drawn from soapy water with a well-defined thickness, but a solid is not necessary to make it. (Likewise, bubbles form at the liquid-air interface in a glass of beer.) Hence, a liquid-vapor interface has the same ability as a solid to entrain matter, providing that surfactants are present at this interface. This is due to the fact that in such a process the surfactants at the surface are not homogeneous in concentration, but diluted by the motion: A gradient of surfactant concentration exists there, and thus a gradient of surface tension, which induces a force that is able to entrain liquid below the surface. If a solid is drawn out of a bath containing surfactants, the film thickness can be expected to be about twice as large as is predicted by the LLD theory for pure liquids, because two interfaces (solid-liquid and liquid-vapor) are now responsible for the coating, instead of one. This is indeed what the experiments show (although more sophisticated arguments involving, for example, the kinetics of adsorption of the surfactant must be taken into account for a complete quantitative understanding of the data). *See* SURFACE COATING.

Coating threshold velocity. A solid slowly drawn out of a liquid which only partially wets it can come out dry, as is commonly observed. For example, a drop moving on a water lily leaves the leaf dry after passing on it. At rest, the liquid meets the solid with a nonzero contact angle. This angle decreases when the solid is moved, but it can remain nonzero thanks to the action of the liquid surface tension which opposes the deformation of the meniscus. Above a threshold in velocity (typically of the order of 1 cm/s or 0.4 in./s), viscous forces dominate capillary ones: A film is deposited and the LLD regime is recovered. The critical velocity of deposition is not yet fully understood, but it has been shown to depend on both the static contact angle and the liquid viscosity.

Self-running droplets. In some cases, the coating may form by itself, instead of being forced to form by an external motion. The so-called running droplets discovered separately by Colin Bain and Thierry Ondarçuhu are a remarkable example (**Fig. 3***a*). These drops contain molecular species which have a double property: first, they are likely to react with the solid; second, the solvent ceases to wet the solid once it is coated with these molecules. Depositing such a drop on a solid generally produces a spontaneous motion (self-running droplets). If the chemical reaction with the solid starts for some reason on one side of the drop, then the two sides sit on different solids: one on the bare solid, the other on the solid coated with the reactive species. This asymmetry provokes the motion, until the reservoir of reactives is exhausted or the whole solid is coated.

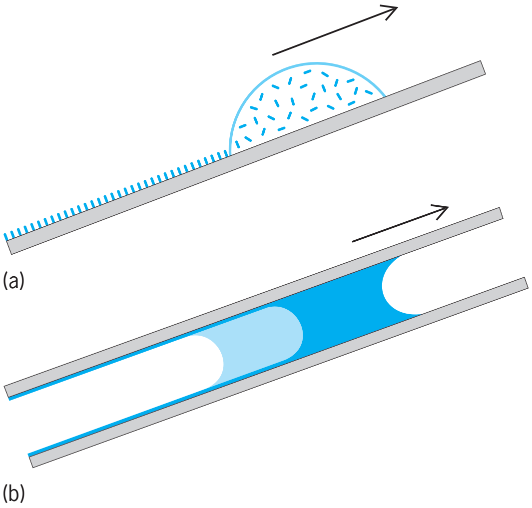


Fig. 3. Self-running droplets. (*a*) Bain-Ondarçuhu device. (*b*) Bico bi-drop.

In 2005, Kenichi Yoshikawa and coworkers proposed reversing the principle: On a glass surface coated with surfactants (and immersed in soapy water), an oil drop is deposited that contains a reactive species that is likely to extract the surfactant. As the oil drop starts taking the surfactant off, it begins to move (again due to a gradient of wettability below the drop). Owing to the redeposition of the surfactant on the glass, the solid surface is regenerated once the drop has moved, allowing the drop to pass the same location many times, until the reactant in the drop is fully saturated with surfactants.

While the Bain-Ondarçuhu drops realize a dry coating of the surface, other possibilities exist if a wet coating is desired. For example, the so-called bi-drops discovered by José Bico, which juxtapose two (wetting) drops inside a tube, also produce a self-motion, again because of the asymmetry of the device (Fig. 3*b*). Each drop can leave a film behind, so that a great variety of coatings (including double coatings) can be realized. The coatings can also be used to achieve better transport properties: The film left behind by the first drop can lubricate the way for the second one. Liquids of very high viscosities can be self-transported by such a device.

In this case, it is in a sense the coating which makes the motion (instead of the motion, the coating, as is usually the case). The driving forces are capillary in essence, so that these droplets can creep upward, as in capillary rise. Research on self-running drops is active and promising, in particular in the context of microfluidics. More sophisticated devices may be possible, where the coating can be addressed and tuned.

David Quéré

Bibliography. B. V. Derjaguin and S. M. Levi, *Film Coating Theory*, 1964; S. F. Kistler and P. M. Schweizer (eds.), *Liquid Film Coating*, 1997; D. Quéré, Fluid coating on a fiber, *Annu. Rev. Fluid Mech.*, 31:347–384, 1999; K. J. Ruschak, Coating flows, *Annu. Rev. Fluid Mech.*, 17:65–89, 1985.

Fluid coupling

A device for transmitting rotation between shafts by means of the acceleration and deceleration of a hydraulic fluid. Structurally, a fluid coupling consists of an impeller on the input or driving shaft and a runner on the output or driven shaft. The two contain the fluid (**Fig. 1**). Impeller and runner are bladed rotors, the impeller acting as a pump and the runner reacting as a turbine. Basically, the impeller accelerates the fluid from near its axis, at which the tangential component of absolute velocity is low, to near its periphery, at which the tangential component of absolute velocity is high (**Fig. 2**). This increase in velocity represents an increase in kinetic energy. The fluid mass emerges at high velocity from the impeller, impinges on the runner blades, gives up its energy, and leaves the runner at low velocity (**Fig. 3**). *See* FLUID MECHANICS; HYDRAULICS.

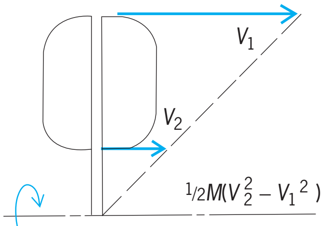


Fig. 2. Hydrokinematic principle.

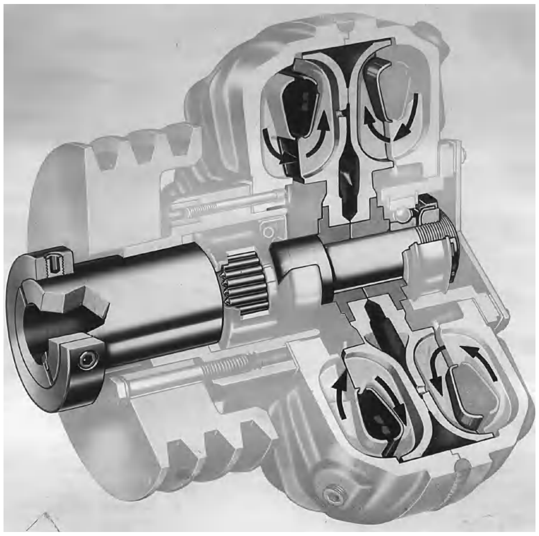


Fig. 3. Sheave with fluid coupling. Dark portion of the figure is driven directly by motor; light portion drives the load. (*Twin Disc Inc.*)

Coupling characteristics. A fluid coupling absorbs power from an engine by means of the centrifugal pump or impeller, which increases the moment of momentum of the mass of fluid. As a consequence, the input power required to drive the coupling increases as the cube of the speed and as the fifth power of the outside diameter of the impeller. Similarly, the capacity of the coupling to absorb torque varies as the square of its input speed.

Because the efficiency of a coupling is a function of slip, couplings normally operate at 95–98% efficiency. The fluid follows a continuous toroidal helix, the pitch of which depends on the slip, or relative difference in rotational velocity, between impeller and runner. Some slip is necessary for energy transfer. As the slip increases at a given input speed, the output shaft turning slower than the input shaft, the quantity of fluid pumped between impeller and runner increases, with consequent increase in torque absorption and transmission. However, because the fluid flow cannot change abruptly, a fluid coupling absorbs rather than transmits shock loads from output to input. Torsional vibrations are similarly absorbed.

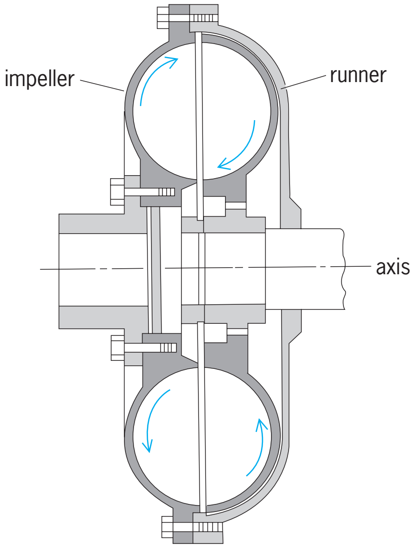


Fig. 1. Basic fluid coupling.

A fluid coupling is designed for a normal running load at 2–5% slip. Conservation of momentum requires that, although shaft speeds may differ, input and output torques must be equal. Thus, for example, in starting a machine from standstill, with the output from the fluid coupling locked, the engine or motor is completely unloaded initially. As the engine starts up toward running speed, the drag torque characteristic of the fluid coupling lies below the developed torque characteristic of the engine. The engine can reach running speed faster with consequent saving in fuel or electrical consumption. The slip of the fluid coupling enables the engine to operate near its maximum-torque speed while accelerating the load, although this characteristic differs from

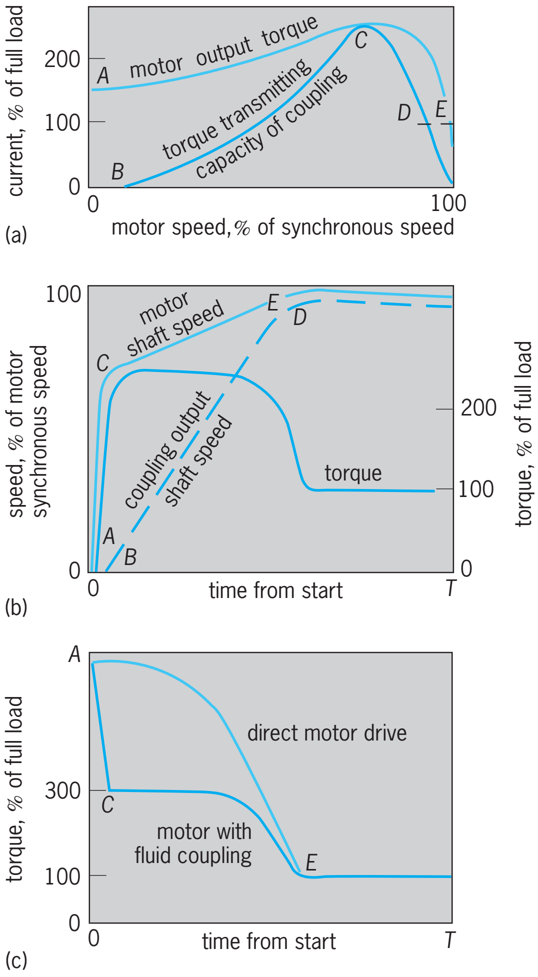


Fig. 4. Starting characteristics of electric motor with fluid coupling. (*a*) Torque speed during acceleration. (*b*) Acceleration. (c) Current.

that of a variable-torque transmission. *See* TORQUE CONVERTER.

Typical application. The effect of a fluid coupling can be described in terms of a specific application. If an electric motor is starting with a high inertia load that requires nearly full-motor torque, the motor accelerates slowly when connected directly to its load. However, with fluid coupling, the motor accelerates rapidly along *A* to *C* in **Fig. 4**. Meanwhile, the torque-transmitting capacity of the coupling builds up along *B* to *C* until it equals the developed motor torque. The motor then can deliver well over full load torque to accelerate the load; the torque demand falls from *C* to *D* as the load comes up to speed *D*. Meanwhile, the motor further accelerates *C* to *E* until steady state is reached with coupling slip represented by differential speed *D* to *E*. Use of fluid coupling improves performance by faster load acceleration and results in appreciably less power loss due to motor heating. *See* AUTOMOTIVE TRANSMISSION; COUPLING.

Henry J. Wirry

Bibliography. W. Bober and R. A. Kenyon, *Fluid Mechanics,* 1980; W. H. Crouse, *Automotive Transmissions and Power Trains*, 5th ed., 1976.

Fluid-flow principles

The fundamental principles that govern the motion of fluids. This article is concerned with these principles and the manner in which they are applied to problems of practical interest. It is assumed that a fluid can be modeled as a continuum and that the details of the molecular structure will appear only in the coefficients associated with constitutive equations or equations of state.

Fundamentals

Understanding the motion of a continuum and the laws that govern that motion is not easy; thus, it should not be surprising that the understanding of motion has its origins in the study of particles, or rigid bodies. While Newton's laws provided a basis for the study of the motion of particles, or more precisely mass points, it was L. Euler (1707–1783) who organized newtonian mechanics into a form which allows a variety of mechanical problems to be solved with relative ease. *See* CLASSICAL MECHANICS; NEWTON'S LAWS OF MOTION.

Euler's laws of mechanics. Euler represented Newton's ideas in terms of two laws of mechanics, which can be stated as follows: (1) The time rate of change of the linear momentum of a body equals the force acting on the body. (2) The time rate of change of the angular momentum of a body equals the torque acting on the body.

It is understood that the distances and velocities associated with these two laws are measured relative to an inertial frame and that the torque and angular momentum are measured relative to the same fixed point. Euler's first law contains all of Newton's three laws except for the part of the third law requiring that the actions of two bodies be directed to contrary parts. This is most often referred to as the central force law, and it is valid only for nonrelativistic systems. *See* ANGULAR MOMENTUM; MOMENTUM; RIGID-BODY DYNAMICS.

In addition to providing a concise statement of the laws of mechanics for continua, Euler postulated that these two laws hold not only for distinct bodies but also for any arbitrary body one might wish to consider. This idea is referred to as the Euler cut principle, and it provides the key to the derivation of governing differential equations from axiomatic statements.

Conservation of mass. The body referred to in Euler's laws consists of some fixed quantity of material that may move and deform in an arbitrary but continuous manner with time t. The region occupied by a body is designated by $V_m(t)$ [**Fig. 1**], and the principle of conservation of mass for a body is stated as follows: The time rate of change of the mass of a body is zero.

Density. If the mass density of the continuum is designated by ρ, the mass of the body is expressed by Eq. (1), where the right-hand side is a volume

$$\text{Mass of a body} = \int_{V_m(t)} \rho \, dV \qquad (1)$$

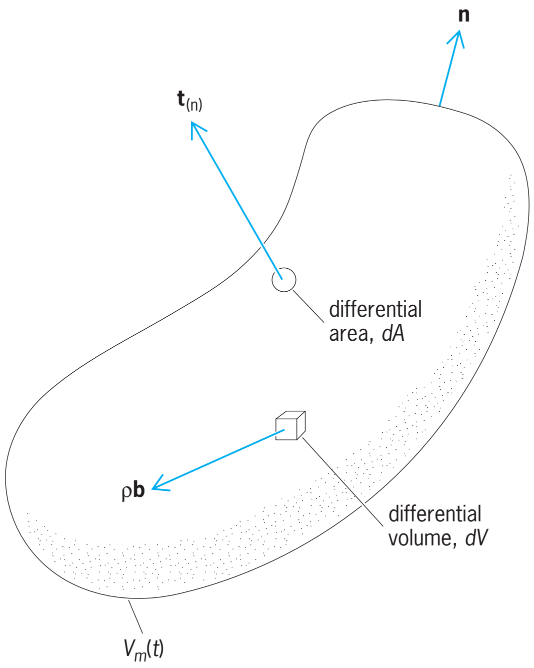


Fig. 1. A material volume.

integral over $V_m(t)$, dV being a differential volume element. The principle of conservation of mass takes the form of Eq. (2), where d/dt represents differentiation with respect to time, or rate of change.

$$\frac{d}{dt}\int_{V_m(t)} \rho \, dV = 0 \tag{2}$$

The density has units of mass/volume, and is nearly constant for liquids and solids. For gases, ρ depends on the temperature and pressure, and an equation of state is required in order to predict the density. *See* DENSITY; GAS; LIQUID.

To obtain a governing differential equation from the axiomatic statement given by Eq. (2), it is necessary to interchange differentiation and integration. G. W. Leibnitz (1646–1716) derived a theorem for accomplishing this, and when this theorem is applied to the time-dependent volume $V_m(t)$, it is usually known as the Reynolds transport theorem. Application of this theorem to Eq. (2) leads to Eq. (3),

$$\int_{V_m(t)} \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right) dV = 0 \tag{3}$$

and since the volume $V_m(t)$ is arbitrary, the integrand must be zero and continuity equation (4) results,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{4}$$

where the first term on the left-hand side represents the time rate of change of mass per unit volume, and the second term represents the net flux of mass, leaving a differential volume. In going from Eq. (3) to Eq. (4), it is assumed that the density ρ and the velocity $\mathbf{v}$ are continuously differentiable with respect to both time and space and that $V_m(t)$ is arbitrary. The velocity vector can be expressed in terms of base vectors and components as Eq. (5), and the gradient operator can be expressed in similar form as Eq. (6). Continuity equation (4) will be used repeatedly below. *See* CALCULUS OF VECTORS.

$$\mathbf{v} = \mathbf{i}v_x + \mathbf{j}v_y + \mathbf{k}v_z \tag{5}$$

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{6}$$

The principle of conservation of mass, as given by Eq. (2), should have obvious intuitive appeal; however, when this concept is applied to a vanishingly small volume (that is, when dV in Fig. 1 is allowed to approach zero), Eq. (4) is obtained and the physical meaning is less clear. When Eq. (4) is integrated over the control volume shown in **Fig. 2**, the macroscopic mass balance is obtained, and for steady flow in a tube it leads to the statement (7).

$$\begin{matrix}\text{Rate at which}\\ \text{mass flows into}\\ \text{the control volume}\end{matrix} = \begin{matrix}\text{Rate at which}\\ \text{mass flows out of}\\ \text{the control volume}\end{matrix} \tag{7}$$

Linear momentum principle. Euler's laws may be expressed in precise mathematical form. The linear momentum of the differential volume element shown in Fig. 1 is the mass ρdV times the velocity $\mathbf{v}$, and this leads to Eq. (8). The total force acting on a body consists of body forces (gravitational, electrostatic, and electromagnetic) that act at a distance upon the mass, and surface forces that act on the bounding surface of the body. As indicated in Fig. 1, the body force per unit volume is represented by $\rho\mathbf{b}$, and the body force acting on the entire body is given by Eq. (9).

$$\begin{matrix}\text{Linear momentum}\\ \text{of a body}\end{matrix} = \int_{V_m(t)} \rho \mathbf{v} \, dV \tag{8}$$

$$\begin{matrix}\text{Body force acting}\\ \text{on the body}\end{matrix} = \int_{V_m(t)} \rho \mathbf{b} \, dV \tag{9}$$

The surface force per unit area is expressed by the stress vector, which is indicated in Fig. 1 by $\mathbf{t}_{(\mathbf{n})}$. The subscript $(\mathbf{n})$ is used as a reminder that the stress vector is a function of the unit outwardly directed normal vector for the surface under consideration. In terms of $\mathbf{t}_{(\mathbf{n})}$, the surface force acting on the body

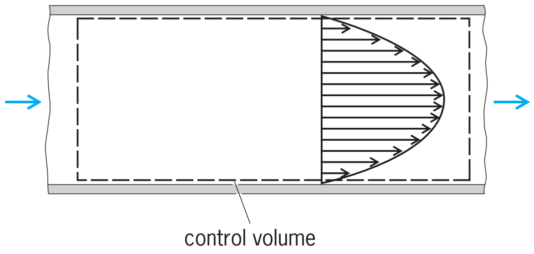


Fig. 2. Steady laminar flow in a tube.

becomes Eq. (10), in which $A_m(t)$ represents the surface of the body illustrated in Fig. 1.

$$\text{Surface force acting on the body} = \int_{A_m(t)} \mathbf{t}_{(\mathbf{n})}\, dA \tag{10}$$

Equations (8) through (10) can be used to express Euler's first law as Eq. (11), where the left-hand side is

$$\frac{d}{dt}\int_{V_m(t)} \rho\mathbf{v}\, dV = \int_{V_m(t)} \rho\mathbf{b}\, dV + \int_{A_m(t)} \mathbf{t}_{(\mathbf{n})}\, dA \tag{11}$$

the time rate of change of momentum, and the terms on the right-hand side represent the body force and the surface force respectively.

The Euler cut principle forced the scientists of the eighteenth century to contemplate the internal state of stress that appeared when an arbitrary body was cut out of a distinct body. It was nearly a century after Euler presented his ideas on the mechanics of continua that A. L. Cauchy (1789–1857) determined the functional dependence of the stress vector on the unit normal vector $\mathbf{n}$. This is given in terms of two lemmas that can be expressed as Eqs. (12) and (13), and both of these relations are a natural consequence of the form of Eq. (11).

$$\mathbf{t}_{(\mathbf{n})} = -\mathbf{t}_{(-\mathbf{n})} \tag{12}$$

$$\mathbf{t}_{(\mathbf{n})} = \mathbf{T}\cdot\mathbf{n} \tag{13}$$

The first lemma is reminiscent of Newton's third law of action and reaction. The second lemma is not so obvious, since it introduces the doubly directed quantity known as the stress tensor. It is best to think of $\mathbf{T}$ as a function that maps the unit normal vector $\mathbf{n}$ into the stress vector $\mathbf{t}_{(\mathbf{n})}$. This idea is shown in matrix form as Eq. (14).

$$\begin{pmatrix} t_{(\mathbf{n})x} \\ t_{(\mathbf{n})y} \\ t_{(\mathbf{n})z} \end{pmatrix} = \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix} \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \tag{14}$$

See MATRIX THEORY.

Angular momentum principle. A precise mathematical statement of Euler's second law is obtained by following the procedure given by Eqs. (8) through (11) to obtain Eq. (15), where the left-hand side represents

$$\frac{d}{dt}\int_{V_m(t)} \mathbf{r}\times\rho\mathbf{v}\, dV = \int_{V_m(t)} \mathbf{r}\times\rho\mathbf{b}\, dV + \int_{A_m(t)} \mathbf{r}\times\mathbf{t}_{(\mathbf{n})}\, dA \tag{15}$$

the time rate of change of angular momentum, and the two terms on the right-hand side represent the torque owing to the body force and the torque owing to the surface force. The mathematical analysis of this equation is extremely complex, and after a great deal of effort it can be proved that Eq. (15) yields only the result that the stress tensor is symmetric. This result can be expressed as Eqs. (16) and indicates that the stress matrix represented in Eq. (14) is symmetric.

$$\begin{aligned} T_{xy} &= T_{yx} \\ T_{xz} &= T_{zx} \\ T_{yz} &= T_{zy} \end{aligned} \tag{16}$$

With this result from Euler's second law, (11) can be used to obtain the differential form of Euler's first law. This is given by Eq. (17), where the

$$\frac{\partial}{\partial t}(\rho\mathbf{v}) + \nabla\cdot(\rho\mathbf{v}\mathbf{v}) = \rho\mathbf{b} + \nabla\cdot\mathbf{T} \tag{17}$$

two terms on the left-hand side represent the time rate of change of momentum per unit volume, and the net flux of momentum leaving a differential volume; and the two terms on the right-hand side represent the body force per unit volume and the surface force per unit volume. Equation (17) is referred to as Cauchy's first equation or the stress equation of motion, since the surface forces are represented in terms of the stress tensor $\mathbf{T}$. The continuity equation in the form of Eq. (18) can be used in order to simplify Eq. (17) to Eq. (19). It is important to think of

$$\mathbf{v}\left(\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v})\right) = 0 \tag{18}$$

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = \rho\mathbf{b} + \nabla\cdot\mathbf{T} \tag{19}$$

Eq. (4) as the governing differential equation for the density ρ and Eq. (19) as the governing differential equation for the velocity $\mathbf{v}$. Euler's second law, given by Eq. (15), provides only a constraint on the stress field, so that there are only six distinct components of the stress tensor $\mathbf{T}$. At this point there are more unknowns than equations, and in order to solve fluid-mechanical problems a constitutive equation for $\mathbf{T}$ is needed.

Fluid characteristics. A fluid is a material that deforms continuously under the action of a shear stress. Air and water are obvious examples of fluids, and their mechanical characteristics are clearly much simpler than catsup, meringue, or toothpaste. It is convenient to separate fluids into those which possess elastic characteristics (often called viscoelastic fluids) and those which do not (often called viscous fluids). In this article, only viscous fluids will be discussed, and for such fluids it is convenient to decompose the stress tensor according to Eq. (20). Here p

$$\mathbf{T} = -p\mathbf{I} + \tau \tag{20}$$

is the thermodynamic pressure, $\mathbf{I}$ is the unit tensor, and t is the symmetric viscous stress tensor. Use of Eq. (20) in Eq. (19) leads to Eq. (21). Here a constitutive equation for t is needed in order to obtain an equation that can be used to determine the velocity field. *See* FLUIDS.

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p + \rho\mathbf{b} + \nabla\cdot\tau \tag{21}$$

Viscosity. The viscosity of a fluid is a measure of its resistance to deformation, and a convenient method of deforming a fluid is illustrated in **Fig. 3**. The lower plate is fixed while the upper plate is moving at a constant velocity u_0, giving rise to a simple shear flow. Experiments of this type have shown that for a large

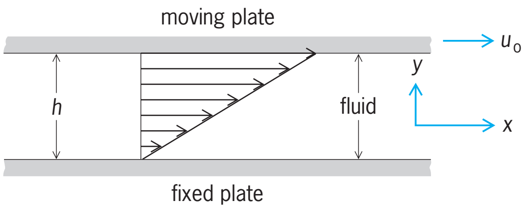


Fig. 3. Velocity distribution between two parallel plates.

class of fluids the force required to maintain the motion of the upper plate is given by Eq. (22). Here F

$$F = \mu\left(\frac{u_0}{h}\right)A \tag{22}$$

represents the force, A is the area of the surface of the plate in contact with the fluid, h is the perpendicular distance between the plates, and μ is the coefficient of shear viscosity. More commonly, μ is referred to simply as the viscosity. If the force per unit area acting in the x direction on a surface having a normal vector parallel to the y direction is designated τ_{yx}, the constitutive equation that is consistent with Eq. (22) is given by Eq. (23). Fluids which obey this relation

$$\tau_{yx} = \left(\frac{F}{A}\right) = \mu\left(\frac{\partial v_x}{\partial y}\right) \tag{23}$$

between the shear stress τ_{yx} and the rate of deformation $\partial v_x/\partial y$ are called newtonian fluids. Many fluids exhibit this type of behavior; however, other fluids do not. These other fluids can be separated into two groups: viscoelastic fluids and non-newtonian fluids. The behavior of non-newtonian fluids is illustrated in **Fig. 4**. The Bingham plastic is representative of fluids which remain rigid until a certain critical shear stress is applied. Such fluids tend to slip at solid surfaces, and in tubes they flow as a plug of toothpastelike material. Shear-thinning fluids tend to become less viscous as they are sheared more vigorously, while shear-thickening fluids become more viscous. A concentrated suspension of cornstarch in water is an example of a shear-thickening fluid. *See* NEWTONIAN FLUID; NON-NEWTONIAN FLUID; VISCOSITY.

Newton's law of viscosity. For complex flows, all six of the distinct components of the viscous stress tensor must be represented in terms of the components of the velocity gradient tensor. For linear, isotropic, viscous fluids, this representation takes the form of Eq. (24). Known as Newton's law of viscosity, it is one

$$\tau = \mu(\nabla\mathbf{v} + \nabla\mathbf{v}^T) + [(\kappa - {}^2\!/_3\mu)\nabla\cdot\mathbf{v}]\mathbf{I} \tag{24}$$

of a series of constitutive equations used to describe the behavior of linear, isotropic materials. Others are Ohm's law of electrical conductivity, Hooke's law of elasticity, and Fourier's law of heat conduction. In Eq. (24), κ is known as the bulk coefficient of viscosity, and it is of importance only for a special class of compressible flows, such as the damping of sonic disturbances. For a very wide class of flows, the second term in Eq. (24) is negligible, and use of this result in Eq. (21) leads to Eq. (25). Here ∇^2 represents the

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p + \rho\mathbf{b} + \mu\nabla^2\mathbf{v} + \mu\nabla(\nabla\cdot\mathbf{v}) + \nabla\mu\cdot(\nabla\mathbf{v} + \nabla\mathbf{v}^T) \tag{25}$$

laplacian operator, and in rectangular coordinates it takes the form of Eq. (26). *See* LAPLACIAN.

$$\nabla^2\mathbf{v} = \frac{\partial^2\mathbf{v}}{\partial x^2} + \frac{\partial^2\mathbf{v}}{\partial y^2} + \frac{\partial^2\mathbf{v}}{\partial z^2} \tag{26}$$

Navier-Stokes equations. Since the viscosity is a function of temperature and pressure, gradients in the viscosity do exist, and $\nabla\mu$ is nonzero. However, the last term in Eq. (25) is usually small compared to the dominant viscous term $\mu\nabla^2\mathbf{v}$, and can be discarded for all but the most sophisticated analyses. In addition, the term $\mu\nabla(\nabla\cdot\mathbf{v})$ is comparable in magnitude to those terms which result from the second term on the right-hand side of Eq. (24) and can therefore be discarded. This leads to Eq. (27), where the

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p + \rho\mathbf{b} + \mu\nabla^2\mathbf{v} \tag{27}$$

two terms within the parenthesis represent the local acceleration and the convective acceleration, and the three terms on the right-hand side represent the pressure force, body force, and viscous force. These equations are known as the Navier-Stokes equations, and serve as the starting point for the analysis of a wide variety of flows. *See* NAVIER-STOKES EQUATION.

Incompressible flow. Since the density is a function of temperature and pressure, there are gradients in ρ for all flows; however, in many flows the variation in ρ can be neglected. These flows are called incompressible flows, and for such flows, a suitable approximation for the continuity equation given by Eq. (4) is Eq. (28). While there are many flows that

$$\nabla\cdot\mathbf{v} = 0 \tag{28}$$

can be treated as incompressible, that is, the density

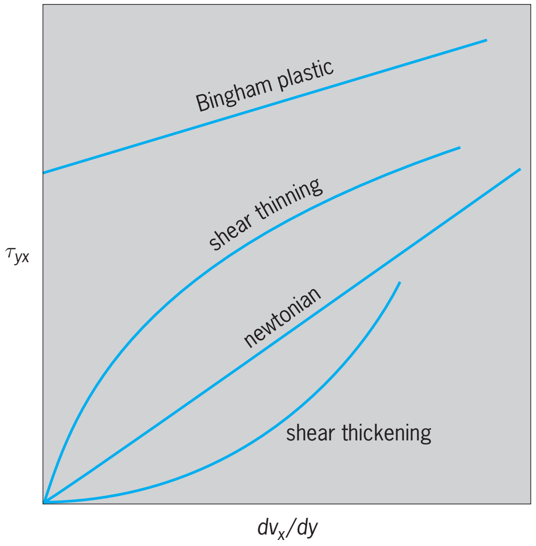


Fig. 4. Behavior of non-newtonian fluids.

can be assigned a constant value, there are no incompressible fluids. Both air and water, for example, are compressible fluids; however, there are many situations for which the flow of both these fluids can be accurately described by Eqs. (27) and (28). In general, this situation occurs when the velocity is small compared to the speed of sound (about 1150 ft/s or 350 m/s for air and about 5200 ft/s or 1600 m/s for water), but there are other situations in which density variations must be considered. The flow that occurs when a pan of water is heated on a stove is an example in which the variation of the density must be taken into account, as is the process of pumping up a bicycle tire. *See* INCOMPRESSIBLE FLOW.

Laminar and turbulent flow. Laminar flow is characterized by the smooth motion of one lamina (or layer) of fluid past another, while turbulent flow is characterized by an irregular and nearly random motion superimposed on the main (or time-averaged) motion of the fluid. The two types of flow can be observed in the heated plume of smoke shown in **Fig. 5**. The smoke originally rises in a smooth laminar stream; however, this buoyancy-driven flow is unstable, and a transition to turbulent flow takes place. This is characterized by an irregular motion as the smoke rises. The transition from laminar to turbulent flow can often be characterized by the Reynolds number, which is a dimensionless group defined by Eq. (29). Here v^* and L^* represent a characteristic velocity and characteristic length respectively. The Reynolds number can be thought of as the ratio of inertial effects to viscous effects, but this is a simplistic interpretation, and a more detailed discussion is below. *See* LAMINAR FLOW; REYNOLDS NUMBER; TURBULENT FLOW.

$$\mathrm{Re} = \frac{\rho v^* L^*}{\mu} \qquad (29)$$

Fig. 5. Laminar and turbulent flow in a smoke plume. (*From S. Whitaker, Introduction to Fluid Mechanics, Krieger Publishing, 1981*)

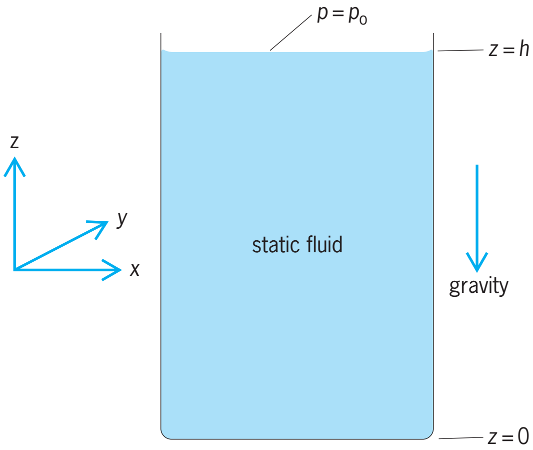


Fig. 6. Pressure variation in a static fluid.

Applications

Some applications of the Navier-Stokes equations to incompressible flows will be considered in order to illustrate the role played by the individual terms in Eq. (27).

Static fluids. When the velocity **v** is zero, Eq. (27) reduces to Eq. (30). If gravity is the only body force under consideration, and the coordinate system shown in **Fig. 6** is used, Eq. (30) reduces to Eq. (31),

$$0 = -\nabla p + \rho \mathbf{b} \qquad (30)$$

$$0 = -\frac{\partial p}{\partial z} - \rho g \qquad (31)$$

where g is the acceleration of gravity. At the surface of the liquid the pressure must be equal to the ambient atmospheric pressure p_0, and under these circumstances Eq. (31) requires that the pressure be given by Eq. (32). Here it has been assumed that the density is constant, and this leads to a linear variation of pressure with position. The situation is not so simple in the atmosphere, since the density of the air depends on both the pressure and the temperature. The hydrostatic variation of the pressure that occurs in the atmosphere is not trivial, and at 10,000 ft (3000 m) above sea level it is responsible

$$p = p_0 + \rho g(h - z) \qquad (32)$$

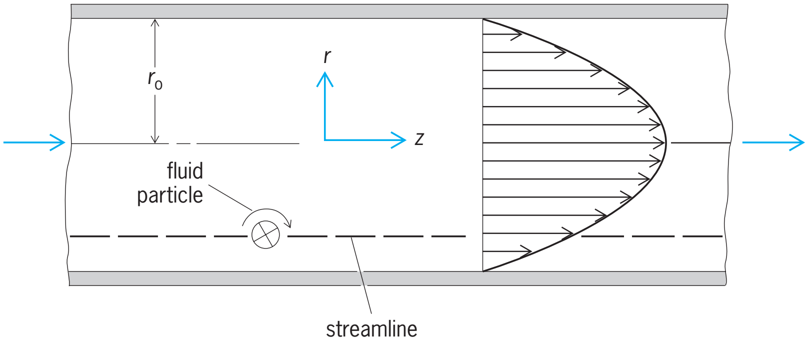


Fig. 7. Uniform flow in a tube.

for a shortness of breath and a lowering of the boiling point of water from 212°F to 193°F (from 100 to 89°C).

Archimedes' principle. The force exerted on an object immersed in a static fluid takes an especially simple form, which is generally attributed to Archimedes: A body is buoyed up by a force equal to the weight of the displaced fluid. The buoyancy force for a constant density fluid is given by Eq. (33),

$$F_{\text{buoy}} = \rho g V \tag{33}$$

where V is the volume of the object and ρ is the density of the fluid. *See* ARCHIMEDES' PRINCIPLE; BUOYANCY; HYDROSTATICS.

Uniform, laminar flow. When steady, laminar flow occurs in a conduit of uniform cross section, the velocity becomes independent of the position along the conduit after a certain distance, referred to as the entrance length. Under these circumstances, Eq. (27) reduces to Eq. (34). For flow in the cylindrical conduit shown in **Fig. 7**, there is only one nonzero component of the velocity vector, and Eq. (34) can be solved to determine the velocity profile, yielding Eq. (35). Here Δp represents the pressure drop that occurs over a distance L, and Eq. (35) indicates that the velocity profile is parabolic, as is illustrated in Fig. 7. If a small amount of dye were injected into a flow of this type, it would trace out a streamline. For this flow, all the streamlines are parallel, and one of these is indicated in Fig. 7. An arbitrarily small element of fluid that moves along this streamline rotates in the manner indicated in Fig. 7. This rotation is caused by the viscous forces, and when viscous effects are small, an irrotational flow is often encountered. When the Reynolds number is greater than 2100, the laminar-flow field becomes unstable and turbulent flow results. The details of the flow are extremely complex, and the irregular motion that accompanies turbulent flow causes the velocity profile to be nearly flat over the central portion of the tube, with very steep velocity gradients near the wall of the tube. *See* PIPE FLOW.

$$0 = -\nabla p + \rho \mathbf{b} + \mu \nabla^2 \mathbf{v} \tag{34}$$

$$v_z = \left(\frac{\Delta p}{L} + \rho b_z\right) \frac{r_0^2}{4\mu} \left[1 - \left(\frac{r}{r_0}\right)^2\right] \tag{35}$$

For turbulent flow, the relation between the volumetric flow rate Q and the pressure drop Δp must be determined experimentally; however, for laminar flow Eq. (35) can be used to obtain Eq. (36). This result is known as the Hagen-Poiseuille law in honor of G. H. L. Hagen and J. L. M. Poiseuille, who experimentally discovered this result in the early 1800s.

$$Q = \frac{\pi r_0^4}{8\mu} \left(\frac{\Delta p}{L} + \rho b_z\right) \tag{36}$$

Boundary-layer flow. Boundary-layer flow is characterized by a large gradient of the velocity in a direction perpendicular to the direction of the main flow, and a small gradient of the velocity in a direction parallel to the direction of the main flow. This type of flow is traditionally discussed in terms of the flow past a thin flat plate, as illustrated in **Fig. 8**, where u_∞ is the uniform fluid velocity (relative to the plate) at large distances from the plate, and $\delta(x)$ is the thickness of the boundary layer. The practical applications of boundary-layer theory have historical roots in the design of airfoils such as that in **Fig. 9**. The gradients of v_x in a boundary-layer flow are characterized by the inequality (37), and this inequality is not valid in the region near the leading edge of the plate, where v_x undergoes rapid changes in the x direction. *See* AIRFOIL.

$$\frac{\partial v_x}{\partial x} \tag{37}$$

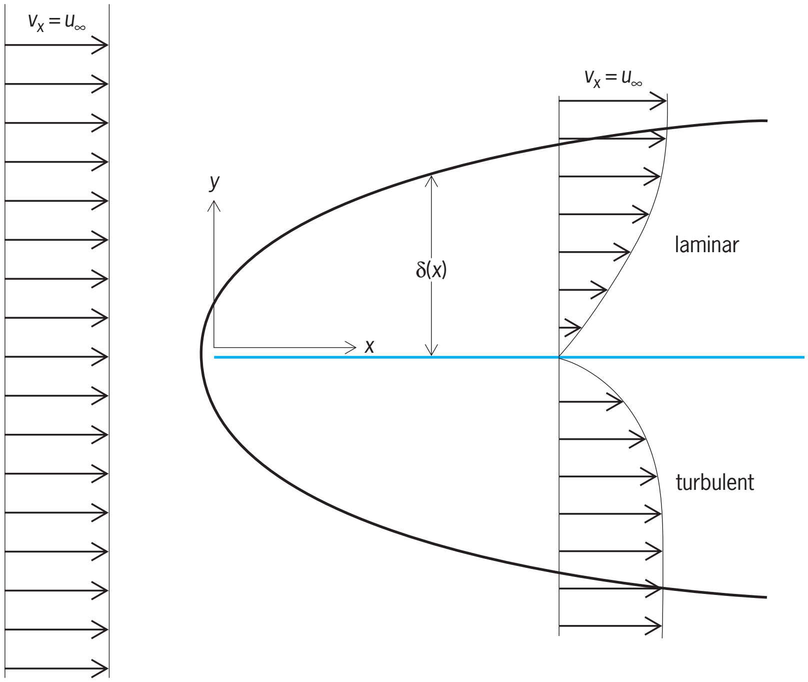


Fig. 8. Boundary-layer flow.

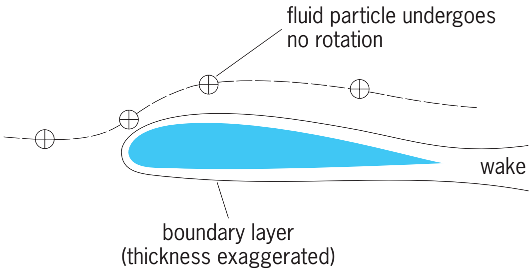


Fig. 9. Diagram of flow past an airfoil.

When the length Reynolds number, given by Eq. (38), is large compared to 1 and the flow is steady,

$$\text{Re}, x = \frac{\rho u_\infty x}{\mu} \tag{38}$$

it is found that the Navier-Stokes equations are characterized by the restrictions of Eq. (39), which indicates

$$0 = -\nabla p + \rho \mathbf{b} \tag{39}$$

that the pressure field is hydrostatic, and inequality (40), which indicates that the flow is quasi–

$$v_x \gg v_y \tag{40}$$

one-dimensional. Under these circumstances, the Navier-Stokes equations and the continuity equation take the form of Eqs. (41) and (42). These two equations

$$\rho\left(v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y}\right) = \mu\left(\frac{\partial^2 v_x}{\partial y^2}\right) \tag{41}$$

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = 0 \tag{42}$$

are usually referred to as the Prandtl boundary-layer equations in the analysis of flow around airfoils. Equations (41) and (42) can be solved to provide values for the velocity field that are in excellent agreement with experimental measurements. In addition, the theoretical values of the velocity can be used to determine the shear stress exerted by the fluid on the flat plate, according to Eq. (43). The drag force acting

$$\tau_{yx}\Big|_{y=0} = \mu \frac{\partial v_x}{\partial y}\Big|_{y=0} \tag{43}$$

on both sides of the flat plate can then be computed from Eq. (44), in which w is the width of the plate.

$$\text{Drag force} = 2w \int_{x=0}^{x=L} \tau_{yx}\Big|_{y=0} dx \tag{44}$$

The drag force determined by Eq. (44) is also in excellent agreement with experimental values, provided the length Reynolds number is constrained by $10^3 \le \text{Re},x \le 10^5$. The reason for this constraint is that Re,x must be larger than 10^3 in order for Eq. (41) to be valid, and for values of Re,x larger than 2×10^5 the flow becomes turbulent. This gives rise to the turbulent velocity profile shown in Fig. 8, and the drag force is significantly larger than that for laminar flow.

While laminar boundary-layer theory applies only to a very limited range of the length Reynolds number, the existence of a thin layer in which viscous effects are important has a profound effect on understanding of fluid-mechanical phenomena. Certainly, the most important of these is the D'Alembert paradox, which results from the misconception that as the velocity becomes progressively larger the inertial term (or convective acceleration) in Eq. (27) should become progressively larger than the viscous term. This line of thought is supported by estimates of the type notation (45), which are based largely on dimensional

$$\rho \mathbf{v} \cdot \nabla \mathbf{v} \simeq \frac{\rho v_2}{\text{length}} \mu \nabla^2 \mathbf{v} \simeq \frac{\mu v}{(\text{length})^2} \tag{45}$$

arguments. If the inertial and viscous length scales are similar (and they need not be), these estimates lead to Eq. (46), and it is tempting to neglect

$$\frac{|\rho \mathbf{v} \cdot \nabla \mathbf{v}|}{|\mu \nabla^2 \mathbf{v}|} = \frac{\rho v\,(\text{length})}{\mu} = \text{Re} \tag{46}$$

the viscous effects in Eq. (27) when the Reynolds number is very large compared to 1. This leads to Euler's equation (47), and under certain circumstances

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v}\right) = -\nabla p + \rho \mathbf{b} \tag{47}$$

this result, along with Eq. (28), can provide a reasonable description of the velocity and pressure fields. However, the most interesting aspect of Eq. (47) is that it can be used to prove that the force exerted by a moving fluid on an airfoil such as the one in Fig. 9 consists of nothing but the buoyancy force. This means that if viscous effects were negligible everywhere in the flow field, the lift force on an airfoil would be zero; and this leads to the (correct) conclusion that an airplane could not fly without the aid of viscosity. Clearly, the boundary layer is of great importance in the analysis of aerodynamic flows. *See* BOUNDARY-LAYER FLOW; D'ALEMBERT'S PARADOX.

Inviscid flow. When the Navier-Stokes equations can be approximated by Euler's equations, the flow is often referred to as inviscid, since one can formally progress from Eq. (27) to Eq. (47) by setting the viscosity equal to zero, that is, $\mu = 0$. From a more realistic point of view, there are two conditions that must be met before Eq. (47) can be considered valid: (1) inertial effects are large compared to viscous effects, which is expressed mathematically by inequality (48), and (2) small causes give rise to small effects.

$$|\rho \mathbf{v} \cdot \nabla \mathbf{v}| \gg \left|\mu \nabla^2 \mathbf{v}\right| \tag{48}$$

A reasonably accurate estimate of the magnitude of the inertial term is given by Eq. (49), in which L_ρ

$$\rho \mathbf{v} \cdot \nabla v = \rho v \frac{d\mathbf{v}}{ds} \simeq \frac{\rho v \Delta v}{L_\rho} \tag{49}$$

is the inertial length scale, and Δv represents the change in the magnitude of the velocity vector that occurs over the distance L_ρ along a streamline. For situations of interest, Δv is often the same order as

υ, and the estimate takes the form of notation (50).

$$\rho \mathbf{v} \cdot \nabla \upsilon \simeq \frac{\rho \upsilon^2}{L_\rho} \tag{50}$$

The magnitude of the viscous term is estimated as in notation (51), in which $\Delta\upsilon$ represents the change in

$$\mu \nabla^2 \upsilon \simeq \frac{\mu \Delta \upsilon}{L_\mu^2} \tag{51}$$

the magnitude of the velocity vector that occurs over the viscous length scale L_μ. Once again, the estimate $\Delta\upsilon \simeq \upsilon$ can be used, and the ratio of inertial effects to viscous effects takes the form of notation (52).

$$\frac{\text{Inertial effects}}{\text{Viscous effects}} \simeq \frac{\rho \upsilon^2 / L_\rho}{\mu \upsilon / L_\mu^2} = \left(\frac{\rho \upsilon L_\mu}{\mu}\right)\left(\frac{L_\mu}{L_\rho}\right) \tag{52}$$

In a laminar boundary layer, $L_\mu = \delta$ and $L_\rho = x$, so that notation (52) takes the form of notation (53).

$$\frac{\text{Inertial effects}}{\text{Viscous effects}} \simeq \left(\frac{\rho u_\infty \delta}{\mu}\right)\frac{\delta}{x} \tag{53}$$

From the Prandtl boundary-layer equations, it follows that viscous and inertial effects are comparable. Thus notation (53) allows an estimate of the boundary-layer thickness as given by notation (54). This is in

$$\delta(x) \simeq \sqrt{\nu x / u_\infty} \tag{54}$$

good agreement with the exact solution obtained from the Prandtl boundary-layer equations, which is given by Eq. (55). In both these results the symbol ν

$$\delta(x) = 4.9\sqrt{\nu x / u_\infty} \tag{55}$$

is used to represent the kinematic viscosity which is given by $\nu = \mu/\rho$.

In order to use notation (52) to decide when viscous effects are negligible, the Reynolds number is defined by Eq. (56) and notation (52) is used to deduce that inertial effects are large compared to viscous effects when inequality (57) is satisfied. When

$$\text{Re} = \frac{\rho \upsilon L_\mu}{\mu} \tag{56}$$

$$\text{Re}\left(\frac{L_\mu}{L_\rho}\right) \gg 1 \tag{57}$$

this inequality is satisfied and small causes give rise to small effects, (47) can be used along with the continuity equation to determine the velocity and pressure fields.

Irrotational or potential flow. When viscous effects are negligible and the velocity and pressure fields are described by Eqs. (28) and (47), the flow is often irrotational. This means that a different element of fluid moves without any rotational motion, such as is illustrated in Fig. 9 for the streamline that is outside the boundary-layer region. When the flow is irrotational, the velocity can be expressed in terms of a potential function according to Eq. (58). From continuity equation (28) it follows that Φ satisfies the laplacian, which can be expressed as in Eq. (59) for rectangular coordinates.

$$\mathbf{v} = -\nabla \Phi \tag{58}$$

$$\nabla^2 \Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0 \tag{59}$$

Solutions of Eq. (59) are relatively easy to obtain for a variety of situations; thus the inviscid-flow approximation allows for the determination of complex flow fields with little mathematical effort. For the flow past an airflow shown in Fig. 9, viscous forces are important in the boundary layer, and the Prandtl boundary-layer equations can be used to analyze the flow. Outside the boundary layer, viscous effects can generally be neglected, and a solution to Eq. (59) provides the velocity field. By matching the two velocity fields at the outer edge of the boundary layer, the complete solution for the flow past an airfoil can be obtained; however, the process is not simple, and it requires considerable mathematical skill and physical insight. *See* LAPLACE'S IRROTATIONAL MOTION.

Bernoulli's equation. When the flow is steady and incompressible, and viscous effects are negligible, a particularly simple form of the Navier-Stokes equations is available. The above restrictions make it possible to write Eq. (27) as Eq. (60). If the body force

$$\rho \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \rho \mathbf{b} \tag{60}$$

$\mathbf{b}$ represents a conservative force field, it can be expressed in terms of the gradient of a scalar according to Eq. (61). When $\mathbf{b}$ represents the gravity vector, ϕ

$$\mathbf{b} = -\nabla \phi \tag{61}$$

is called the gravitational potential function, and for the coordinate system shown in Fig. 6, ϕ is given to within an arbitrary constant by Eq. (62). Since the flow is incompressible, Eq. (63) is valid and Eq. (60) takes the form of Eq. (64). The velocity vector can be

$$\phi = gz + \text{constant} \tag{62}$$

$$\rho \mathbf{b} = -\rho \nabla \phi = -\nabla(\rho \phi) \tag{63}$$

$$\rho \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla(p + \rho \phi) \tag{64}$$

expressed in terms of the magnitude υ and the unit vector $\mathbf{l}$ illustrated in **Fig. 10**, so that the left-hand side of Eq. (64) can be expressed as Eq. (65). Here

$$\rho \mathbf{v} \cdot \nabla \mathbf{v} = \rho \upsilon \lambda \cdot \nabla \mathbf{v} = \rho \upsilon \frac{d\mathbf{v}}{ds} \tag{65}$$

s represents the arc length measured along any one

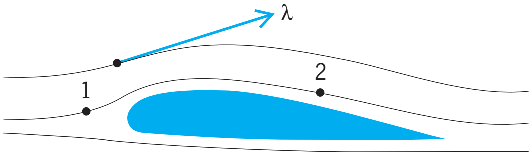


Fig. 10. Streamlines around an airfoil.

of the streamlines shown in Fig. 10, and the definition of the directional derivative given by Eq. (66)

$$\lambda \cdot \nabla = \frac{d}{ds} \qquad (66)$$

has been used. Substitution of Eq. (65) into Eq. (64) yields a form that does not appear to be suitable for problem solving, Eq. (67). However, the scalar product

$$\rho v \frac{d\mathbf{v}}{ds} = -\nabla(p + \rho\phi) \qquad (67)$$

(or dot product) of this equation with the unit tangent vector is Eq. (68), and a little analysis leads

$$\rho v \lambda \cdot \frac{d\mathbf{v}}{ds} = -\lambda \cdot \nabla(p + \rho\phi) \qquad (68)$$

to Eq. (69). This indicates that the quantity ${}^1/_2\rho v^2 + p + \rho\phi$

$$\frac{d}{ds}\left({}^1/_2\rho v^2 + p + \rho\phi\right) = 0 \qquad (69)$$

is a constant along a streamline, and Eq. (62) can be used to express this idea as Eq. (70). This is

$${}^1/_2\rho v^2 + p + \rho g z = C \qquad (70)$$

known as Bernoulli's equation. For the two points identified on a streamline shown in Fig. 10, Eq. (70) can be used to derive Eq. (71). For irrotational flows,

$${}^1/_2\rho v_2^2 + p_2 + \rho g z_2 = {}^1/_2\rho v_1^2 + p_1 + \rho g z_1 \qquad (71)$$

it can be shown that the constant in Eq. (70) is the same for all streamlines, and most flows for which viscous effects are negligible fall into this category. Under these circumstances, Eqs. (58) and (59) are used to determine the velocity field, and Eq. (70) to determine the pressure field. *See* BERNOULLI'S THEOREM; FLUID MECHANICS; GAS DYNAMICS; HYDRODYNAMICS. Stephen Whitaker

Bibliography. H. Aksel and O. C. Eralp, *Gas Dynamics*, 1993; I. G. Currie, *Fundamental Mechanics of Fluids*, 2d ed., 1993; H. Schlichting and J. Kestin, *Boundary Layer Theory*, 8th ed., 2000; W. G. Vincenti and C. H. Kruger, Jr., *Introduction to Physical Gas Dynamics*, 1965, reprint 1975; S. Whitaker, *Fundamental Principles of Heat Transfer*, 1977, reprint 1983; S. Whitaker, *Introduction to Fluid Mechanics*, 1968, reprint 1992.

Fluid mechanics

The engineering science concerned with the relation between the forces acting on fluids (liquids and gases) and their motions, and with the forces caused by fluids on their surroundings. A fluid subjected to unbalanced forces or stresses continues in motion as long as unbalanced forces are applied. For example, in the pouring of water from a pitcher the water velocity is very high over the lip, moderately high approaching the lip, and very low near the bottom of the pitcher. The unbalanced force is gravity, that is, the weight of the tilted water particles near the surface. The flow continues as long as water is available and the pitcher remains tilted.

In contrast, a solid deforms under forces and stresses and simply assumes a new deformed position and remains static. It does not flow.

Laypersons instinctively recognize this distinction between a flowing liquid and a statically deformed solid. In fact, not only liquids but vapors and gases are defined as fluids, because of this characteristic of continuous deformation. A fluid cannot withstand an applied shear stress, however small, without moving. Solids remain static because they resist shear. *See* FLUIDS; STRENGTH OF MATERIALS.

Types of fluids. A fluid may be a liquid, vapor, or gas. In thermodynamics the three states of matter are solid, liquid, and vapor. The term vapor denotes a gaseous substance interacting with its own liquid phase, for example, steam above water. If this phase interaction is not important, the vapor is simply termed a gas. *See* THERMODYNAMIC PROCESSES; VAPOR PRESSURE.

Gases have weak intermolecular forces and expand to fill any container. Left free, gases expand and form the atmosphere of the Earth. Gases are highly compressible; doubling the pressure at constant temperature doubles the density. *See* GAS.

Liquids, in contrast, have strong intermolecular forces and tend to retain constant volume. Placed in a container, a liquid occupies only its own volume and forms a free surface which is at the same pressure as any gas touching it. Liquids are nearly incompressible; doubling the pressure of water at room temperature, for example, increases its density by only 0.005%. *See* OPEN CHANNEL.

Liquids and vapors can flow together as a mixture, such as steam condensing in a pipe flow with cold walls. This constitutes a special branch of fluid mechanics, covering two-phase-flow. *See* CAVITATION.

Density and other properties. The physical properties of a fluid are essential to formulating theories and developing designs for fluid flow. Especially important are pressure, density and temperature.

The density of a fluid is its mass per unit volume, denoted as ρ. The most common units for density are kg/m^3 in the International System (SI) and lb/ft^3 or slugs/ft^3 in the English systems. Liquids are typically hundreds of times more dense than gases; at room temperature and pressure, $\rho = 998$ kg/m^3 (62.4 lb/ft^3 or 1.94 slugs/ft^3) for water and only 1.20 kg/m^3 (0.0075 lb/ft^3 or 0.00233 slug/ft^3) for air. Fluid density is independent of location, whether on the Earth or the Moon.

In practical problems, the specific weight of a fluid or weight per unit volume is often used, denoted as $\gamma = \rho g$, with units of newton per cubic meter (N/m^3; SI) or pound-force per cubic foot (lbf/ft^3; English). The quantity g is the acceleration of gravity, with an average value on Earth of 9.81 m/s^2 (32.2 ft/s^2). Thus, for water at standard conditions, $\gamma = 9790$ N/m^3 (62.4 lbf/ft^3). *See* DENSITY.

The pressure of a fluid, denoted as p, is the normal force per unit area exerted on any small elemental surface in the fluid. Standard atmospheric pressure is 101,325 N/m^2 or pascals (2116 lbf/ft^2). Differences in pressure, or pressure gradients, often drive fluid flows. *See* PRESSURE.

The temperature of a fluid, denoted as T, expresses the average energy of the fluid molecules. It may be stated in convenient units (°F or °C) or in absolute units (K or °R). Standard room temperature is 20°C (293 K) or 68°F (528°R). *See* CARBOHYDRATE; TEMPERATURE.

Ideal gas. For all fluids, p, ρ, and T are related by the thermodynamic state relation for the substance. For liquids, which are nearly incompressible, ρ is assumed to be constant, independent of p or T. For gases, a convenient approximation is the ideal gas law, given in Eq. (1), where R is the gas constant. For

$$p = \rho RT \tag{1}$$

any particular gas, R is computed by dividing a universal physical constant, 8315 (m^2)/(s^2 · mole · K), by the molecular weight of the gas. For air, with an average molecular weight of 29, $R = 8315/29 = 287$ (m^2)/(s^2 · K). In Eq. (1), T must have absolute units.

The ideal gas law becomes inaccurate at very high temperatures, where chemical reactions may occur, or at lower temperatures where the gas is near its saturated liquid state. The latter case is illustrated by steam, which is near the liquid state and sufficiently nonideal for extensive tables of its (p, ρ, T) relations to have been published for engineers and scientists. *See* STEAM.

Hydrostatic condition. Since shear stresses cause motion in a fluid and result in differences in normal stresses at a point, it follows that a fluid at rest must have zero shear and uniform pressure at a point. This is the hydrostatic condition. The fluid pressure increases deeper in the fluid to balance the increased weight of fluid above. For liquids, and for gases over short vertical distances, the fluid density can be assumed to be constant. Then the pressure increases linearly with depth. If point 2 is vertically below point 1 at distance h, the two pressures are related by hydrostatic formula (2). This formula is illustrated

$$p_2 = p_1 + \rho g h = p_1 + \gamma h \tag{2}$$

in **Fig. 1** for air and water, whose densities must be inserted into Eq. (2). If the liquid volume is limited, it forms a free surface, as in Fig. 1, which is level or horizontal and at uniform pressure.

For gases over large changes in height, such as the atmosphere, Eq. (2) must be modified for varying density and results in a nonlinear formula.

The linearly varying vertical change of hydrostatic pressure in Eq. (2) gives simple results for overall pressure on a plane surface. Integration of (pressure) × (area) over the surface gives Eq. (3) for the

$$F_{\text{plane}} = p_{cg} A_{\text{plane}} \tag{3}$$

force on a plane area A_{plane} of any shape, where p_{cg} denotes the fluid pressure at the centroid of the plane surface, as determined by Eq. (2). Equation (3) is also illustrated in Fig. 1. Formulas such as Eq. (3) are used to design submerged surfaces, including dams and submarine hulls, to withstand large hydrostatic forces.

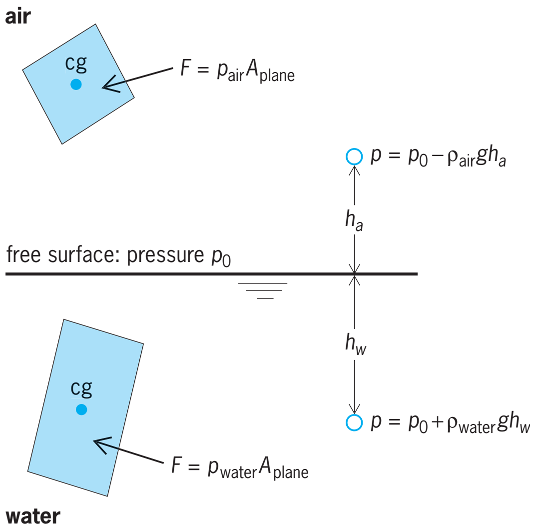


Fig. 1. Hydrostatic fluid properties. cg = centroid of the plane.

When the pressure forces are integrated over a closed three-dimensional surface (an immersed body), Archimedes' two famous laws of buoyancy result: (1) A body immersed in a fluid experiences a vertical buoyant force equal to the weight of the fluid it displaces. (2) A floating body displaces its own weight in the fluid in which it floats. These principles are widely used in applied fluid-flow problems such as the design of ships and underwater vehicles. *See* ARCHIMEDES' PRINCIPLE; HYDROSTATICS.

Viscosity. When a fluid is subjected to shear stress, it flows and resists the shear through molecular momentum transfer. The macroscopic effect of this molecular action, for most common fluids, is the physical property called viscosity. Shear stress results in a gradient in fluid velocity; the converse is also true.

Figure 2 shows a typical case of a fluid flow whose x-directed velocity u varies with transverse distance y, $u = u(y)$. This convenient scheme of showing a row of velocity vectors with arrowheads connected by a continuous curve is called a velocity profile. At any point in the fluid, the slope of the velocity profile is a measure of local shear stress τ. For all gases and vapors and for the common liquids (water, gasoline, oil), slope and shear are proportional as in Eq. (4).

$$\tau = \mu \frac{du}{dy} \tag{4}$$

$$\mu = \mu(p, T)$$

The coefficient μ is called the viscosity of the fluid, with units of kg/(m · s), lb/(ft · s), or slug/(ft · s). The viscosity is a physical property of a fluid and is determined experimentally, for example, by shearing

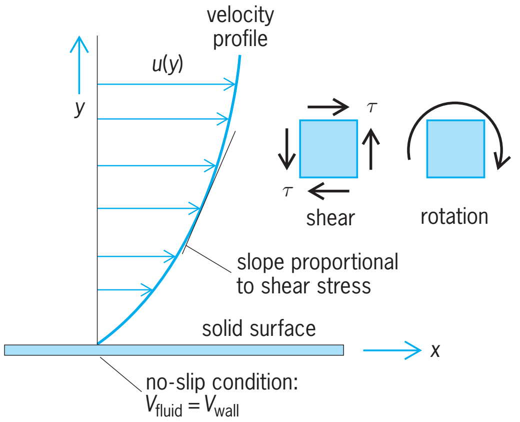


Fig. 2. No-slip condition and shear stress set up by velocity gradients.

the fluid between two plates or two concentric cylinders.

Gases have low viscosity, light liquids such as water and mercury are in the medium range, and oils have very high μ. At 20°C (68°F) and 1 atm (101 kPa), three representative examples of viscosity are air [1.8×10^{-5} kg/(m · s)], water [1.0×10^{-3} kg/(m · s)], and glycerin [1.5 kg/(m · s)]. The viscosity of liquids decreases with temperature, whereas that of gases increases. The very slight increase of fluid viscosity with pressure is usually neglected.

The common fluids for which the linear relationship of Eq. (4) holds are called newtonian viscous fluids. More complex fluids, such as paints, pastes, greases, and slurries, exhibit nonlinear or non-newtonian behavior and require more complex theories to account for their behavior. *See* NEWTONIAN FLUID; NON-NEWTONIAN FLUID; VISCOSITY.

No-slip condition. A common characteristic of all fluids, whether newtonian or not, is that they do not slip at a solid boundary. No matter how fast they flow away from the boundary, fluid particles at a solid surface become entrapped by the surface structure. The macroscopic effect is that the fluid velocity equals the solid velocity at a boundary. This is called the no-slip condition, illustrated in Fig. 2, where the solid is fixed, so that the fluid velocity drops to zero there: $u(0) = 0$. No-slip sets up a slow-moving shear layer or boundary layer when fluid flows near a solid surface, as shown in Fig. 2. The theory of boundary-layer flow is well developed and explains many effects involving viscous flow past immersed bodies or within passages. *See* BOUNDARY-LAYER FLOW.

Kinematic viscosity. Fluid-flow problems often involve the ratio of viscosity to density, called the kinematic viscosity ν given in Eq. (5). The mass units

$$\nu = \frac{\mu}{\rho} \qquad (5)$$

cancel (one meaning of kinematic is "independent of mass") and the units of ν are m^2/s or ft^2/s.

Although gases have low viscosity, they also have low density and their kinematic viscosity is thus often higher than for common liquids. For example, at 20°C (68°F) and 1 atm (101 kPa) air has a viscosity $\nu = 1.51 \times 10^{-5}$ m^2/s, whereas for water $\nu = 1.01 \times 10^{-6}$ m^2/s, or 15 times lower. As a crude estimate, this makes water 15 times more likely than air to undergo low-viscosity effects such as transition to turbulence.

Compressibility and speed of sound. All fluids are at least slightly compressible, that is, their density increases as pressure is applied. In many flows, however, compressibility effects may be neglected.

Fluid compressibility can be characterized by the isentropic bulk modulus K and the speed of sound a. The bulk modulus, given in Eq. (6), is the change

$$K = -v\left(\frac{\delta p}{\delta v_s}\right) = \rho\left(\frac{\delta p}{\delta \rho_s}\right) \qquad (6)$$

in pressure for a fractional change in fluid density for an isentropic process, where v is the specific volume and s the fluid entropy. A large value of K indicates a fluid which is difficult to compress. For example, at 20°C and 1 atm (101 kPa), $K = 2.19 \times 10^9$ N/m^2 (4.57×10^7 lbf/ft^3) for water and only 1.42×10^5 N/m^2 (2.97×10^3 lbf/ft^2) for air.

The speed of sound a is defined as the velocity of propagation of a small pressure disturbance through the medium. Application of conservation of mass, momentum, and energy to such a disturbance results in Eq. (7). Using data given above for K and ρ at 20°C

$$a = \left(\frac{K}{\rho}\right)^{1/2} = \left(\frac{\delta p}{\delta \rho}\right)_s^{1/2} \qquad (7)$$

and 1 atm (101 kPa) gives $a \simeq 1480$ m/s (4860 ft/s) for water and $a \simeq 340$ m/s (1117 ft/s) for air. For an ideal gas, Eq. (7) may be written, with the help of Eq. (1), as Eq. (8), where $k = c_p/c_v$ is the ratio of

$$a = (kRT)^{1/2} \qquad (8)$$

specific heats and is approximately 1.4 for air. *See* SOUND.

A very important parameter in determining compressibility effects is the Mach number, or ratio of flow velocity V to fluid speed of sound, given by Eq. (9). For subsonic flow, $\text{Ma} < 1$, whereas for su-

$$\text{Ma} = \frac{V}{a} \qquad (9)$$

personic flow, $\text{Ma} > 1$. The flow is essentially incompressible if $\text{Ma} < 0.3$; hence for air the flow velocity is less than about 100 m/s (330 ft/s). Almost all liquid flows and many gas flows are thus treated as incompressible. Even a supersonic airplane lands and takes off in the incompressible regime. *See* MACH NUMBER.

A problem involving compressibility of liquids arises when a valve is suddenly opened or closed in a piping system. Strong pressure waves may be propagated through the pipes and cause serious damage. *See* COMPRESSIBLE FLOW; WATER HAMMER.

Ideal incompressible flow. Flow analysis can be greatly simplified by neglecting friction and compressibility and assuming that the flow is steady or varying only in space. The resulting equations of motion of ideal flow may be solved analytically for many problems. The velocity field **V**(x, y, z) may be plotted in various ways to reveal the flow pattern, and the pressure field $p(x, y, z)$ is directly related to velocity and local elevation by Bernoulli's equation for incompressible, frictionless, steady flow, Eq. (10), where c

$$\frac{p}{\rho} + \frac{V^2}{2} + gz = c \qquad (10)$$

is a constant along a streamline and z is the vertical height of any point in the flow. A streamline is defined as a line everywhere parallel to the local velocity vector. If velocity and pressure are known, the force may be computed on any boundary or internal surface in the flow.

Equation (10) is a weak form of the first law of thermodynamics, balancing the work of pressure forces with changes in kinetic and potential energy. Bernoulli's equation, though widely used, is thus inaccurate if there are significant heat-transfer effects, shaft work, or viscous losses. *See* BERNOULLI'S THEOREM.

Flow visualization. Fluid flow is highly visual, and many photographs and sketchs of flow patterns have been published. Movies and video tapes have been made showing flow phenomena. Flow patterns are shown as streamlines, which are everywhere tangent to the local velocity vector; pathlines, which trace the trajectory of a single particle; and streaklines, which are the locus of particles issuing from a single point. Streamlines arise naturally in the mathematical theory of flow, whereas pathlines and streaklines arise in experimental studies using smoke, bubbles, dyes, or neutral particles. In steady fluid flow the three types of lines are identical.

Figure 3 shows the streamline pattern for ideal frictionless flow past a circular cylinder. The lines provide a qualitative feeling for the flow. Where they are equally spaced at the far left and far right, the velocity is uniform. Where they are widely spaced near the left and right of the cylinder, the velocity is low and the pressure high. Where they are closely packed above and below the cylinder, the velocity is high and the pressure low. Figure 3 has some inaccuracies due to viscous effects.

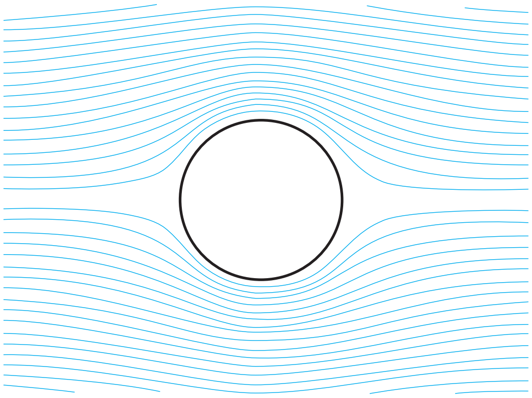


Fig. 3. Streamlines for uniform flow of a frictionless, incompressible fluid past a circular cylinder. This idealized flow pattern exhibits double symmetry.

Viscous flow effects. The assumption of frictionless flow is an idealization which may be inapplicable to the actual flow of a real or viscous fluid. In particular, where solid walls are present, the no-slip condition, as in Fig. 2, slows down the fluid near the wall and creates a boundary layer where shear effects are important. This viscous layer generally grows in thickness as it moves downstream along the wall. It may ultimately encompass a large portion of the downstream flow pattern.

In the case of flow into a pipe or duct, the shear layers grow and meet in the pipe center and thereafter fill the pipe with viscous effects. Bernoulli relation (10) is readily modified for pipe flow by adding a friction-loss term. If points 1 and 2 are upstream and downstream in the pipe flow, respectively, the modified Bernoulli relation is given by Eq. (11), where h_f

$$\frac{p_1}{\rho} + \frac{V_1^2}{2} + gz_1 = \frac{p_2}{\rho} + \frac{V_2^2}{2} + gz_2 + gh_f \qquad (11)$$

is a loss term expressed in length units (m or ft). Equation (11) shows that the total flow energy decreases downstream because of friction losses. The loss term may be correlated with flow velocity, pipe size, and physical properties of the fluid. Equation (11) may not be used if there are significant heat-transfer or shaft-work effects.

In the case of flow past immersed bodies, as in Fig. 3, viscous effects are variable and involve more than just a loss term. The flow pattern is strongly dependent upon the dimensionless parameter called the Reynolds number, given by Eq. (12), where V is

$$\mathrm{Re}_D = \frac{\rho V D}{\mu} \qquad (12)$$

the stream velocity and D a suitable length scale, in this case the cylinder diameter. *See* REYNOLDS NUMBER.

Figure 4 shows the real-fluid flow past a circular cylinder at a Reynolds number of 10^5. The boundary layer on the front of the body is thin, and outside this layer the flow is approximately frictionless and satisfies Bernoulli relation (10). However, as it rounds the shoulder of the body, the low-velocity boundary layer is unable to resist the rising pressure toward the rear of the body. It separates from the body and forms a broad, pulsating wake which produces a large downstream force on the body, called drag. This is a sharp contrast to the ideal theory of Fig. 3, which predicts zero drag and was thought in the nineteenth century to be a zero paradox, although actually it is simply the incorrect result of neglecting viscous effects. *See* D'ALEMBERT'S PARADOX; WAKE FLOW.

Figure 4 is a striking example of the importance of viscous effects, but other flows are not so strongly modified. Ideal-flow theory has been used to predict such phenomena as airfoil and cascade forces, flow through orifices, and surface-wave motions, and is a useful first approximation in flow analysis.

Laminar and turbulent flow. An interesting aspect of fully viscous flow within the boundary layer or inside a long duct is the existence of two types of flow.

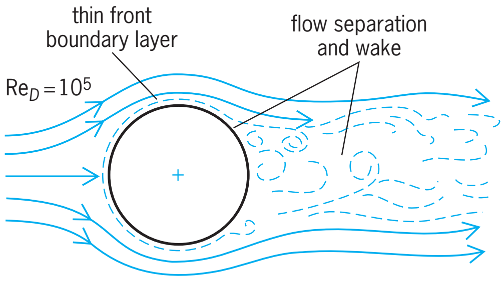


Fig. 4. Actual flow pattern for viscous flow past a circular cylinder at a moderately high Reynolds number. The outer stream is grossly perturbed by broad flow separation and wake. (*After F. M. White, Fluid Mechanics, 2d ed., McGraw-Hill, 1986*)

Although flow instrumentation was relatively crude, it was known in the early nineteenth century that flow could be either steady and orderly or fluctuating and disorderly. In 1883, O. Reynolds discovered that the controlling parameter is the Reynolds number from Eq. (12). By introducing a dye streak into a pipe flow, Reynolds was able to visually demonstrate the two flow regimes.

Figure 5 shows Reynold's sketches of the two types of flow. At low Re_D, the dye streak remains straight and smooth (Fig. 5*a*), as if the fluid particles remained in distinct laminae which merely sheared each other; this regime is now called laminar flow. At higher Re_D, the dye streak seems to grow and burst and fill the pipe with color (Fig. 5*b*). Reynold's spark photograph of this condition (Fig. 5*c*) reveals multiple eddies and fast-changing whorls in the flow; this second regime is now called turbulent flow.

Numerous pipe-flow experiments verify that the flow is laminar until Re_D reaches a critical or transition value of about 2000. The transition range, intermittently laminar and turbulent, is from 2000 to 4000. For $Re_D > 4000$, the flow becomes fully turbulent.

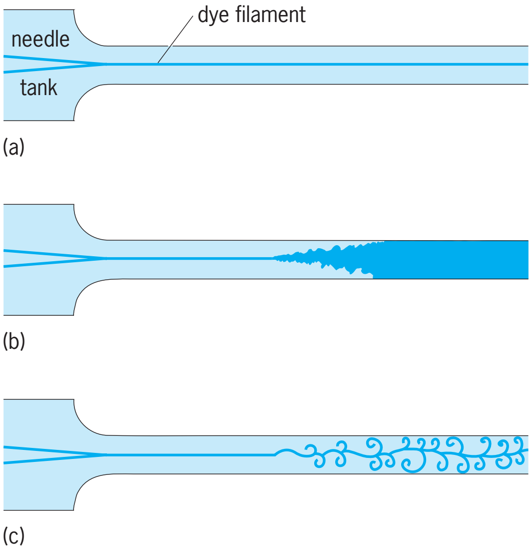


Fig. 5. Pipe-flow transition experiment. (*a*) Low Re_D, laminar flow. (*b*) High Re_D, turbulent flow. (*c*) Spark photograph of turbulent flow condition. (*After O. Reynolds, an experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous and of the law of resistance in parallel channels, Phil. Trans. Roy. Soc., London, A174:935–982, 1883*)

This same laminar-transition-turbulence sequence occurs for all types of viscous flows. The critical Reynolds number depends upon the geometry and the presence of disturbances which might trigger turbulence, such as rough walls, noisy pumps, loudspeakers, and heat-transfer effects. *See* LAMINAR FLOW; TURBULENT FLOW.

Pipe flow. Flow in a pipe is a good example of laminar versus turbulent behavior. Since the 1940s, instrumentation has been developed, such as hot wires, hot films, tiny flush-mounted piezometers, and laser-Doppler meters, which measure all the details of fluctuating turbulent velocities, temperatures, and pressures in a flow. *See* FLOW MEASUREMENT.

One of the most striking differences between laminar and turbulent flow is the shape of the mean velocity profiles (**Fig. 6**). The laminar profile (Fig. 6*a*) is a paraboloid shape which can be predicted mathematically. Its mean value V is exactly one-half its maximum value $u_{\max}$, and the friction loss may be exactly predicted by Eq. (13) from Eq. (11). In contrast,

$$h_{f,\text{laminar}} = \frac{32\mu LV}{\rho g D^2} \tag{13}$$

the turbulent profile (Fig. 6*b*) is flat, with extremely steep slopes at the wall. Its average velocity varies from 81 to 89% of the maximum velocity, depending upon Re_D. Its friction loss also varies with Re_D several times larger than suggested by Eq. (13), and is predicted not by theory but by empirical correlations. *See* PIPE FLOW.

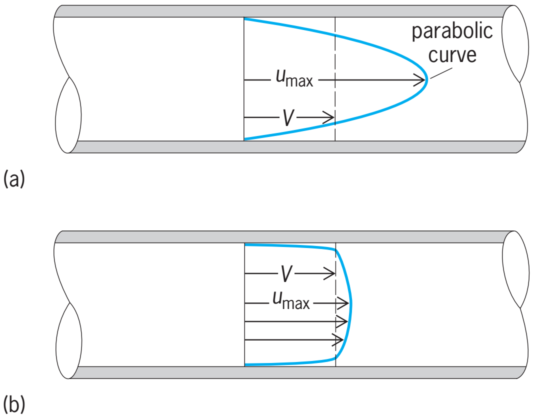


Fig. 6. Comparison of (*a*) laminar and (*b*) turbulent pipe flows at the same volume flow rate (m^3/s). V = average flow velocity; $u_{\max}$ = maximum flow velocity. (*After F. M. White, Fluid Mechanics, 2d ed., McGraw-Hill, 1986*)

Generally speaking, laminar flows can be predicted analytically and modeled accurately with digital computer simulations, whereas turbulent flows require empirical correlations and other approximate modeling tools, which make turbulence predictions, though improving with research, uncertain and lacking in generality.

Compressible flow. As the Mach number of a flow increases beyond 0.3, compressibility becomes important. The density begins to change significantly and is coupled to temperature and pressure changes. Mass and momentum conservation must be combined with the laws of thermodynamics and a state equation for the gas.

Energy relation (11) must be modified to account for internal energy changes, heat transfer, and shaft and viscous work. The result is given in Eq. (14).

$$h_1 + \frac{V_1^2}{2} + gz_1 = h_2 + \frac{V_2^2}{2} + gz_2 - q + w_s + w_v \qquad (14)$$

Here, q is the heat per unit mass transferred to the fluid, and w_s and w_v are the shaft work and viscous work, respectively, per unit mass done by the fluid, and h is the fluid enthalpy per unit mass, defined by Eq. (15), where e is the internal energy per unit

$$h = e + \frac{p}{\rho} \qquad (15)$$

mass. *See* ENTHALPY; HEAT; INTERNAL ENERGY.

The most common case is compressible flow in a duct, where viscous work is zero because of the no-slip condition. Assuming adiabatic flow, no shaft work, and negligible height changes, Eq. (14) reduces to Eq. (16), where c is a constant.

$$h_1 + \frac{V_1^2}{2} = h_2 + \frac{V_2^2}{2} = c \qquad (16)$$

See ISENTROPIC FLOW.

Equation (16) does not require frictionless flow. If Eq. (16) is combined with the ideal gas law, Eq. (1), and the definition of the Mach number, Eq. (9), the result in Eq. (17) relates temperature at any point to

$$T = \frac{T_0}{1 + 0.5(k-1)\mathrm{Ma}^2} \qquad (17)$$

the stagnation temperature T_0 at a point where the velocity is zero.

If frictionless flow is assumed, the motion is at constant entropy and the local pressure may be related to stagnation pressure, p_0, as in Eq. (18). Thus, in adi-

$$p = \frac{p_0}{\left[1 + 0.5\,(k-1)\,\mathrm{Ma}^2\right]^{k/(k-1)}} \qquad (18)$$

abatic frictionless duct flow, both the fluid temperature and pressure decrease significantly with Mach number.

A converging-diverging duct is shown in **Fig. 7**. For steady one-dimensional flow, the mass conservation relation is given in Eq. (19), where A is the area

$$\rho_1 A_1 V_1 = \rho_2 A_2 V_2 = \text{constant} \qquad (19)$$

of the duct. If the duct is fed by a source pressure p_0 and the receiver pressure p_e is gradually reduced, the throat pressure falls until the throat Mach number reaches unity, from which, for $k = 1.4$, Eq. (18) predicts that $p = 0.53p_0$. The throat is sonic, and any further reduction in p_e fails to change the flow rate, because signals can no longer be propagated upstream from the throat. The nozzle is said to be choked, and the flow downstream from the throat becomes supersonic. This interesting behavior is nearly impossible in liquid flow but is easily accomplished in gases by moderate changes in boundary pressures. Other aspects of compressible flow are the formation of shock waves and the behavior of vehicles moving faster than sound. *See* GAS DYNAMICS; NOZZLE; SHOCK WAVE; SUPERSONIC FLIGHT.

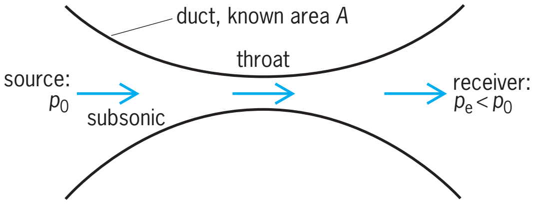


Fig. 7. Compressible duct flow through a converging-diverging nozzle. As receiver pressure drops, the throat becomes sonic and the downstream flow supersonic.

Boundary-layer theory. As shown in Fig. 2, the no-slip condition creates a slow-moving boundary layer when a flow is near a solid surface. Both shear and rotation (local fluid angular velocity) within the boundary layer are large at the surface and gradually decrease to zero in the outer, nearly inviscid, stream. The boundary layer is very thin if the Reynolds number is above 10^4. A similar thermal boundary layer exists if the wall is hot or cold compared to the stream flow.

Boundary layers are laminar near the leading edge of the surface and undergo transition to turbulence as they grow thicker downstream. The theory is well developed and predicts both laminar and turbulent velocities as well as temperatures plus wall shear stress and heat transfer. The theory also predicts flow separation on the rear (or rising-pressure) region of bodies, as in Fig. 4. After the separation point, boundary-layer theory is generally no longer accurate. *See* FLUID-FLOW PRINCIPLES; HYDRODYNAMICS; RAREFIED GAS FLOW.

Frank M. White

Bibliography. R. W. Fox, A. T. McDonald, and P. J. Pritchard, *Introduction to Fluid Mechanics*, 6th ed., 2003; *Illustrated Experiments in Fluid, Mechanics*, NCFMF Book of Film Notes, 1972; I. H. Shames, *Mechanics of Fluids*, 4th ed., 2003; V. L. Streeter, E. S. Wylie, and K. W. Bedford, *Fluid Mechanics*, 9th ed., 1998; M. Van Dyke, *An Album of Fluid Motion*, 1982; F. M. White, *Fluid Mechanics*, 5th ed., 2003.

Fluidization

The processing technique employing a suspension or fluidization of small solid particles in a vertically rising stream of fluid—usually gas—so that fluid and solid come into intimate contact. This is a tool with many applications in the petroleum and chemical process industries. There are major possibilities of even greater growth and utility in the future. Suspensions of solid particles by vertically rising liquid streams are of lesser interest in modern processing, but have been shown to be of use, particularly in liquid contacting of ion-exchange resins. However, they come in this same classification and their use involves techniques of liquid settling, both free and hindered (sedimentation), classification, and density flotation. *See* ION EXCHANGE; MECHANICAL CLASSIFICATION.

Properties of fluidized systems. The interrelations of hydromechanics, heat transfer, and mass transfer in the gas-fluidized bed involve a very large number of factors which must be understood for mathematical and physical analysis and process design. Because of the excellent contacting under these conditions, numerous chemical reactions are also possible—either between solid and gas, two fluidized solids with each other or with the gas, or most important, one or more gases in a mixture with the solid as a catalyst. In the usual case, the practical applications in multimillion-dollar plants have far outrun the exact understanding of the physical, and often chemical, interplay of variables within the minute ranges of each of the small particles and the surrounding gas phase. Accordingly, much of the continuing development of processes and plants is of an empirical nature, rather than being based on exact and calculated design.

The fluidized bed results when a fluid, usually a gas, flows upward through a bed of suitably sized, solid particles at a velocity sufficiently high to buoy the particles, to overcome the influence of gravity, and to impart to them an appearance of great turbulence, similar to that of a violently boiling liquid. Fluid velocities must be intermediate between that which would lift the particles to maintain a uniform suspension and that which would sweep the particles out of the container. The fluidized bed is in a relatively stable condition of vigorous contacting of fluid and solids, with a lower boundary at the point of fluid inlet and an upper definite and clearly marked boundary surface at which the gas disengages itself from the system. The upper boundary usually goes up or down with increase or decrease of the gas velocity.

Particle sizes are often in the range of 30–125 micrometers, but they may vary to $^1/_4$ in. (6 mm) or more. Superficial gas velocities may range 0.02–1.0 ft/s (0.006–0.3 m/s), depending on factors such as the relative densities of the gas and solid, the size and shape of solid particles, and the number of particles per cubic foot (bed density) desired.

Conditions within the bed are intermediate between that of a packed column and that of a pneumatic transporter of the particles. In pneumatic transport, the slippage of the gas past the solid particles is often relatively small, and the components of friction and inertia carry the solid particle along. In the fluidized bed, the slippage of gas past the solid particle is sufficient to balance the gravitational effect or weight, but is not sufficient to supply a continuing velocity.

An increase in fluid velocity increases the lift or upward drag on the particles. This causes them to rise, thereby increasing the bed voidage and the effective cross section available for gas passage and also decreasing the number of particles per unit volume and the density of the bed as a whole. This decreases the interstitial velocity to reduce the drag on the individual particles, which then settle somewhat until the gravitation and the fluid-dynamic forces on the particles ultimately come into balance. Further increase in fluid velocity causes further bed expansion until the bed passes beyond the limits of the containing vessel and the particles are transported.

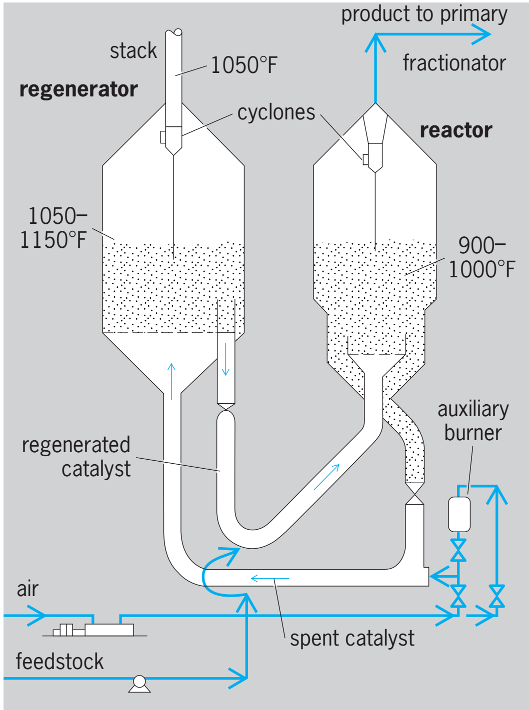
There is excellent contact between the solid particles themselves and the particles with the gas. If there is a temperature difference, there is an excellent flow of heat between the particles themselves and between the particles and the gas. Particularly if heat is absorbed, as in drying of a solid by a hot gas, or if heat is given off because of chemical reaction, heat transfers to equate the temperatures of the gas and the solid very quickly.

Similarly, there is a major opportunity for mass transfer by molecular diffusion to or from the solid particles and between the solid particles and the gas phase. Examples are the movement of reactants in the gaseous phase toward the surface of solid particulate catalysts, and the reverse movement of products from the catalytic surface back to the body of the gas phase.

Principal applications. With such excellent opportunities for heat and mass transfer to or from solids and fluids, fluidization has become a major tool in such fields as drying, roasting, and other processes involving chemical decomposition of solid particles by heat. An important application has been in the catalysis of gas reactions, wherein the excellent opportunity of heat transfer and mass transfer between the catalytic surface and the gas stream gives performance unequaled by any other system. In the petroleum industry, for example, the cracking of hydrocarbon vapors in the presence of solid catalysts affords many advantages. Among these are a minimum tendency for overheating of one particle or one part of the catalytic bed as compared to another, and the immediate removal of the reactant products, resulting in greatly increased efficiency of the catalyst. The catalyst itself often becomes carbonized on the surface because of the coking tendency of the reacting hydrocarbons. This catalyst, as shown in the **illustration**, may then be removed by pneumatic transport or otherwise to a second fluidized reactor wherein an oxidation is accomplished, again with maintenance of relatively uniform conditions. The rejuvenated catalyst may be returned to the prime reactor by pneumatic or other transfer lines, and the process can be continued indefinitely.

Some fluid reaction systems have been standardized in design, permitting the operation of units comparable in maximum size to other types of units in the adjacent steps of processing.

From a process-development standpoint, the dense bed of particles and gas means that enough



Catalytic cracking process using fluidized catalyst. °C = (°F − 32)/1.8.

catalyst can be maintained in a relatively small vessel to give a high degree of conversion with a low pressure drop causing gas flow. The mechanical design of a fluidization system and accessories is comparable, in many respects, to that of a body of liquid. However, special devices are needed to control fluid and solid handling, to introduce gas to the fluid bed, and to control the level of the solid particles, which usually are allowed to overflow the top of vertical pipes, as in maintaining the level of a true liquid. Vapor-disengaging units, filters, and cyclones, for removal of minute amounts of fines which might otherwise escape, blowers, pumps, pneumatic transport lines, and bucket elevators are also needed.

One important component in many units, besides the fluidizing chamber—often the reactor—is the standpipe, which allows a recycling of solids and builds up pressure on the solids for transfer purposes. In a large catalytic cracking unit, the rate of circulation of solids between vessels may be as much as 50 tons/min (0.75 metric ton/s).

Fluidized versus fixed beds. While catalytic particle-gas reactions have long been done in fixed beds, these cannot all be converted to fluidized operations. There are notable advantages and disadvantages of each. The advantages of the fluidized bed have been evidenced in the much better mixing, which results in greatly improved heat transfer, mass transfer, and uniformity. Thus there may often be better heat and material recovery, that is, increased yields, lower catalyst losses, and sometimes lower pressure drops. Fluidized beds allow ready additions of solids, combustion products, or direct combustion within the bed to give a much greater flexibility of process design and operation.

On the other hand, the concurrent flow in a fluidized reactor is disadvantageous, and multibed reactors are necessary to secure the advantages of countercurrent flow. There may be mixing of products with raw materials throughout the bed, and particularly when the solid is a reactant, it is difficult to obtain a high conversion and almost impossible to obtain a 100% conversion, except in a multibed system. Some attrition and thus loss of catalyst particles and changing fluidization properties with changing average size result. Recovery equipment for fines may be necessary, either because of their value or their nuisance as air pollutants if lost to the atmosphere. Similarly, there is considerable erosion of the walls, and especially of transfer lines. Because of the nature of fluidization, many reactions requiring a catalyst bed cannot be so operated, for example, those that give waxy or gummy materials which would agglomerate on the catalyst particles and also those wherein properties of the solids or of the gas stream, the necessary gas velocities, and so on, are outside of the mechanical possibilities of the fluidization operation. Thus, fixed beds are in many cases either required or most advantageous.

Individual industries. The simplest use of a fluidized bed is as a simple heat exchanger, with direct contact between solids and fluid—either liquid or gas—possibly to recover heat from a hot gas or from a hot solid by contacting with a cold solid or a cold gas, respectively. Such a heat exchanger, with no chemical reactions, is often used as a supplementary fluidized bed to one wherein there is a chemical reaction. Within a single bed, there is no possible countercurrent action, although multiple beds can be used and their particle flow arranged countercurrently to that of the fluid.

A basic example of diffusion is in the use of particles of an adsorbent, such as activated carbon, to remove vapors of solvent from a fluidizing stream of carrier gas, such as air. The air loses the vapor molecules by their diffusion first through a gas film surrounding the particles and then into the solid mass of the particles. The solvent-charged particles overflow the top of the bed to form a second bed, fluidized usually with steam to desorb the solvent. The steam-solvent mixture is separated by distillation, and the solvent-free particles are recycled to the first bed. Obviously, this example of mass transfer of molecules by diffusion to and from the particles also includes a transfer of heat in both beds.

A similar bed without chemical reaction is used as a dryer of solid particles. In this example, the moisture must diffuse to the surface of the particles, from which it is evaporated into the gas stream. It must again be noted that, because the flows in the bed are not countercurrent, a single unit cannot achieve complete dryness, except with an excess in heat consumption.

Very similar is the removal of chemically combined water and carbon dioxide, or both, in what is usually called calcination, at temperatures about 1400–2200°F (760–1200°C). Multiple beds in countercurrent are used to minimize heat requirement. A necessary added expense is the grinding of limestone to a size which is practical to fluidize, usually not over $^1/_4$ in. (6 mm) maximum dimension. The crushed limestone descends from a top-drying fluidized bed through a series of beds to one where combustion of fuel oil is also taking place to supply the heat for calcination both in this bed and in the higher beds. The lowest bed receives and cools the burnt lime with the incoming air for combustion, which is thus preheated.

Roasting of pyrites to give sulfur dioxide is practiced throughout the world in possibly hundreds of fluidized beds; the combustion operation may be controlled better than in other types of roasting furnaces. Similarly, additional roasters are used to convert the iron oxide from hematite to magnetite, which then is heated to give pellets for steel manufacture.

Other sulfide ores are roasted in closely controlled fluidized beds to give metal sulfates, which may then be leached to recover their metal values from the gangue.

There are other examples of fluidized beds where there are more complicated chemical reactions between gases and solid. Included are the carbonization of wood and of coal to form charcoal and coke with concomitant volatile products, the combustion of pulverized coal and other solid fuels, and the production of portland cement clinker. More modern applications have been the production of titanium chloride in a bed of a ground titanium mineral, rutile, and coke fluidized with chlorine and the production of uranium compounds for nuclear fuels.

Fluid coking is an interesting cracking process, in which the undistillable residues of petroleum towers are coked in a fluidized reactor at about 1000°F (540°C) to give coke and a distillate. Coke is deposited on the fluidized particles of seed coke, prepared by fine grinding of some of the product, and these particles are built up to the desired product sizes for withdrawal and sale.

The greatest use in number and size of fluidized reactors and in volume of throughput has been in catalytic cracking and reforming of petroleum and its fractions. Scores of variations have developed the processing by particles of catalyst petroleum vapor streams. The modern petroleum industry could not otherwise operate under present conditions and costs. *See* ADSORPTION OPERATIONS; BULK-HANDLING MACHINES; CATALYSIS; CRACKING; DIFFUSION; FLUIDIZED-BED COMBUSTION; HEAT TRANSFER; UNIT OPERATIONS.

Donald F. Othmer

Bibliography. J. F. Davidson (ed.), *Fluidization*, 1978; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 10, 1994; M. Kwauk (ed.), *Advances in Chemical Engineering*, vol. 20: *Fast Fluidization*, 1994; O. Levenspiel and D. Kunnii, *Fluidization Engineering*, 2d ed., 1991.

Fluidized-bed combustion

A method of burning fuel in which the fuel is continually fed into a bed of reactive or inert particles supported by upflowing air, which causes the bed to behave like a turbulent fluid. Fluidized beds have long been used for the combustion of low-quality, difficult-to-burn fuels and more recently have been developed for the clean burning of coal.

Process description. A fluidized-bed combustor is a furnace chamber whose floor is slotted, perforated, or fitted with nozzles to uniformly distribute upflowing air. The chamber is partially filled with maximum-sized particles which fluidize at the airflow rate required for combustion. When fluidization takes place, the bed of material expands (bulk density decreases) and exhibits the properties of a liquid. As air velocity increases, the particles mix more violently, and the surface of the bed takes on the appearance of a boiling liquid. If air velocity were increased further, the bed material would be blown away.

Once the temperature of the bed material is raised to the fuel-ignition temperature with start-up burners, fuel can be fed into or over the bed to achieve the desired operating temperature. Proper selection of air velocity, operating temperature, and bed material will cause the bed to function as a chemical reactor. In the case of high-sulfur coal combustion, the bed is generally limestone, which reacts with and

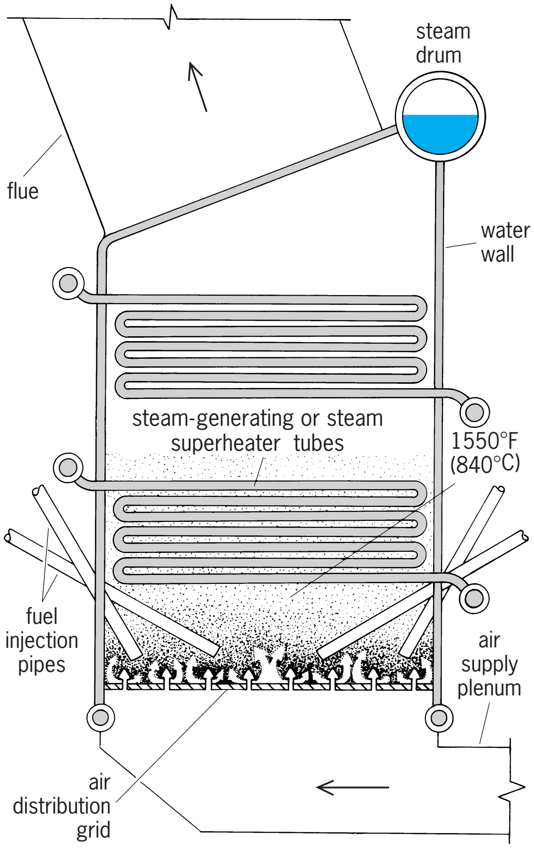


Fig. 1. Fluidized-bed steam generator.

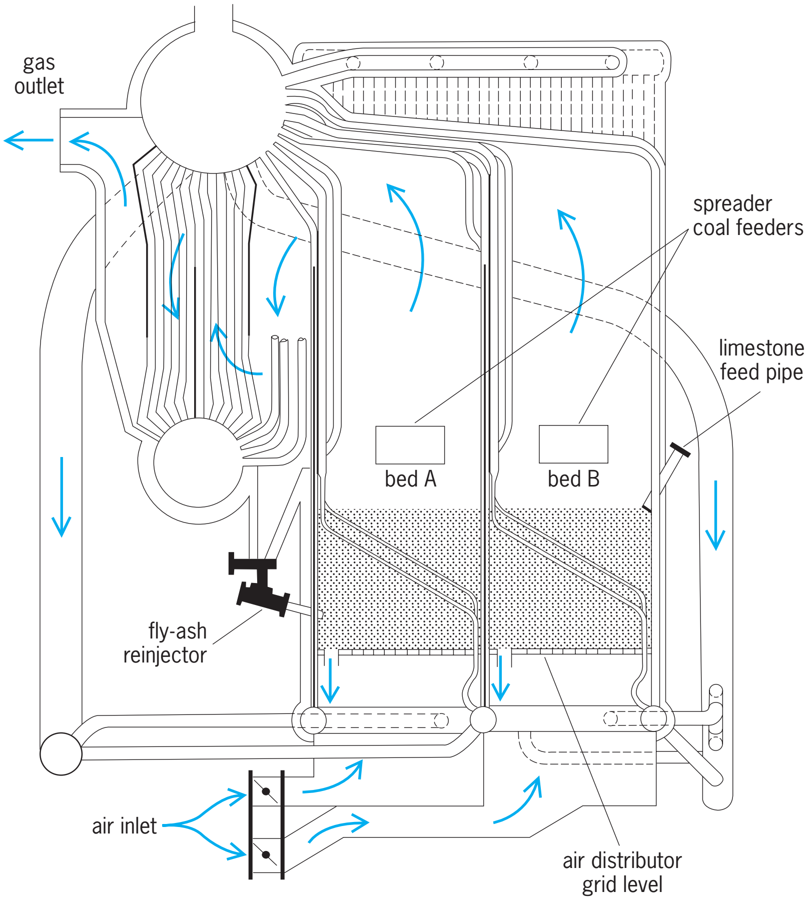


Fig. 2. Flow diagram showing the operation of a multicell fluidized-bed boiler.

absorbs the sulfer released during burning, reducing sulfur dioxide emissions. The alkali compounds in coal from the western United States can react with the sulfur dioxide to limit emissions.

Fluidized-bed combustion can proceed at low temperatures (1400–1500°F or 760–840°C), so that nitrogen oxide formation is inhibited. In the limestone-sulfur–absorbing designs, the waste product is a relatively innocuous solid, although the large volume produced does have environmental implications. Researchers in both federal and privately funded programs are exploring techniques for utilizing the waste product, for example, determining its potential value as an agricultural supplement or bulk aggregate. Low combustion temperatures minimize furnace corrosion problems. At the same time, intimate contact between the hot, turbulent bed particles and the heat-transfer surfaces results in high heat-transfer coefficients. The result is that less surface area is needed, and overall heat-transfer tube requirements and costs are lower. Intensive development of fuel-fired systems has been undertaken, particularly for pressurized operation, where the movement of solids through the pressure barrier poses some special problems. These efforts appear successful in pilot plants.

Although commonly regarded as a sulfur emissions control process, fluidized-bed combustion is applicable to low-sulfur fuels and wastes. In firing a low-sulfer coal from the western United States, for example, the bed material would simply be coal ash.

Applications. There are three areas of application: incineration, gasification, and steam generation.

Coal gasification. The original application of fluidized-bed combustion was to the gasification of coal. A gasifier burns coal in a fluidized bed with less air than that amount which is required for complete combustion. This results in a high concentration of combustible gases in the exhaust. This principle is being applied in a number of pilot plants which, for the most part, are directed at producing low-sulfur-

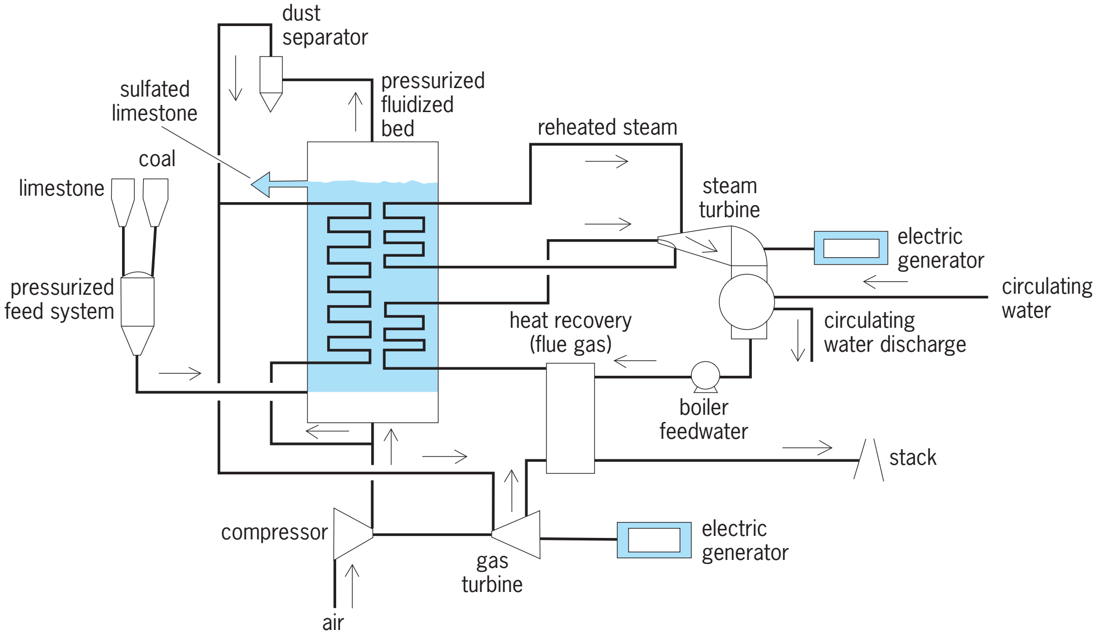


Fig. 3. Flow diagram of 600-MW combined-cycle system.

content gases of medium heating value (150–300 Btu per standard cubic foot or 5590–11,170 kilojoules per cubic meter). *See* COAL GASIFICATION.

Incineration. The use of this technology for efficient incineration of waste products, such as food-processing wastes, kraft process liquors, coke breeze, and lumber wastes, has increased rapidly. The process has the advantage of controllability, lower capital cost, more compact design, and reduced emissions.

Steam generation. Heat-transfer tubes are immersed within the hot fluidized bed, resulting in a compact steam boiler (**Fig. 1**). Combining cells of individual beds permits control of the output of larger-capacity boilers (**Fig. 2**). Beds may be arranged horizontally, vertically, or both ways in the same system.

Fluidized-bed steam generators with capacity as great as 350,000 lb/h (160,000 kg/h) are in operation throughout the world firing a variety of low-quality fuels. In Europe, emphasis has been on cogeneration systems and firing wastes. Industrial boilers firing coal and mining wastes are typical applications of the technology in the United States.

Operation and testing of a 20-megawatt pilot plant and initial steps for a 200-megawatt demonstration plant have resulted from cooperative efforts of the Electric Power Research Institute and the Tennessee Valley Authority.

In the so-called pressurized fluidized-bed combustor, operating the combustion chamber at elevated pressures permits the use of the heated exhaust gases to drive gas turbines. Tests have been conducted on a 13-MW pilot plant using this combined cycle. Designs have been made for larger plants (**Fig. 3**). *See* GAS TURBINE; STEAM-GENERATING UNIT.

Other applications. Fluidized-bed combustion techniques have had a long history of application in the chemical industries, especially in catalytic petroleum cracking, and in extractive metallurgy (ore roasters and calciners). Development of a fluidized-bed combustor for the spent graphite fuel rods from high-temperature gas-cooled nuclear reactors has been undertaken.

Competitiveness of process. Since nearly any combustible product, including such difficult fuels as oil shales, anthracite wastes, and residual oils, can be effectively burned, fluidized-bed combustion has gained worldwide commercial status. Environmentally, its competitiveness with flue-gas desulfurization, coal cleaning, and advanced technologies such as liquefaction has been established. *See* COAL LIQUEFACTION; COMBUSTION; FLUIDIZATION.

Michael Pope

Bibliography. E. P. Johnson, Fluidized-bed combustion: Better fluid-bed units to make debuts, *Chem. Eng.*, p. 39, April 1982; L. N. Rubow (ed.), *Fluidized Bed Combustion, 1993: FBC's Role in the World Energy Mix*, 1993; *Summary Evaluations of Atmospheric Pressure Fluidized Bed Combustion Applied to Electric Utility Large Steam Generators*, EPRI Report FP 308, prepared by Babcock and Wilcox Company, Alliance, Ohio, 1976.

Fluids

Substances that flow under shear stress and have a density range from essentially zero (gases) to solidlike values (liquids). Fluids comprise one of the two major forms of matter. Solids, the other form, generally deform very little when shear forces are applied, and their densities do not change significantly with pressure or temperature.

At the atomic level, the relative molecular positions in solids are well defined to very great distances, and the molecules rarely change locations. However, fluids have a uniform structure beyond a few molecular diameters. **Figure 1** schematically demonstrates the differences in molecular distribution in gases, liquids, and solids according to what x-ray and neutron scattering measurements indicate about the probability of finding molecular centers at any separation. Probabilities different from unity indicate "structure." For all phases and densities, the repulsive intermolecular forces prevent any molecules from approaching to small separations, so the probability at such separations is zero. Just beyond the repulsive barrier, the probability is higher in all phases due to attractive forces. For low-density gases, the fluid is uniform beyond the first neighbors and there is no real structure. For the denser liquids, there is only local structure, whose extent depends on the density. Fluids do not show the long-range structure observed in solids. Also, molecules in fluids change positions and orientations more frequently than do those in solids. In fact, at low densities, there is essentially no inhibition to molecular motion except for the container walls. Fluids have been conceptualized as mostly solidlike, with regions of vacuum or free individual molecules, and there even have been models, based on such ideas, that succeed in describing fluid properties. Fluid structure is not heterogeneous, as this concept implies. *See* CRYSTAL STRUCTURE; INTERMOLECULAR FORCES; NEUTRON DIFFRACTION; SOLID-STATE PHYSICS; X-RAY DIFFRACTION.

The distinction between solids and fluids is most easily seen in substances and mixtures which show a well-defined melting process. For substances with large molecules, such as polymers, ceramics, and biologicals, this distinction is less clear. Instead, there is a slow evolution of structure and of resistance to flow

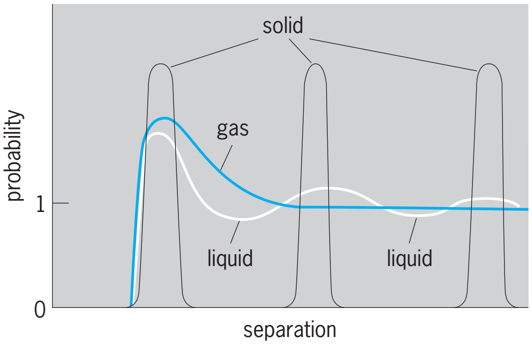


Fig. 1. Probability of finding molecular centers at various separations in gases, liquids, and solids.

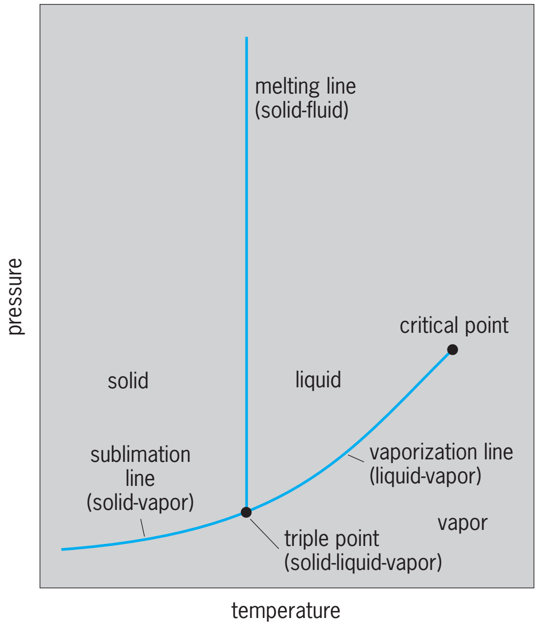


Fig. 2. Conditions of pure-component phase behavior.

as temperature or some other variable is changed. *See* GLASS TRANSITION; MELTING POINT; POLYMER.

Vapors and liquids. Molecular density varies greatly in fluids and is their most important characteristic. The distinction between vapors (or gases) and liquids is most clear for substances and mixtures that show well-defined vaporizing (boiling) and condensing processes. The high-density liquid boils to make a low-density gaseous vapor. **Figure 2** shows the pressures and temperatures for which pure substances are single phases. At the conditions of the lines between the single-phase regions, two phases can be observed to coexist. At the state of intersection of the lines (the triple point), three phases can coexist. For most substances, the triple-point pressure is well below atmospheric. However, for carbon dioxide, it is very high, so that dry ice sublimes rather than melts, as water ice does. Beyond the end of the liquid-vapor (saturation or vapor-pressure) line, vaporization and condensation cannot be observed. The state at the end of this line is called the critical point, where all the properties of the vapor and liquid become the same. There is no such end to the solid-liquid (or melting) line, because the solid and fluid structures cannot become the same. *See* CRITICAL PHENOMENA; PHASE TRANSITIONS; TRIPLE POINT.

Figure 3 shows the relationship of pressure and density of fluids at temperatures above the melting point and at pressures above the triple point. As the pressure on a vapor is increased at fixed temperature (T_1), the density changes rapidly until the condensation point is reached. During condensation, the pressure on the two phases is constant until all the vapor is gone. Then the fluid is relatively incompressible, and the density changes very little with even large changes in pressure. The region of conditions where two phases are observed is capped by the critical point. At a temperature just above the critical (T_2), the variation of density with pressure is very great, even though the system is one phase. At the highest temperatures (T_3), the pressure-density variation is more gradual everywhere. The relation between fluid-state variables of pressure P, density ρ and temperature T is often represented by a quantity called the compressibility factor, $z = P/\rho RT$ where R is a universal constant equal to 1.987 calories/mol K, 1.987 Btu/lb mole °R, or 8.3143 J/mol K and where density is in units of moles per unit volume. The behavior of z is shown in **Fig. 4**. At the lowest pressures and highest temperatures (T_3 and T_4), z approaches 1, where the fluid is called an ideal gas and the molecules behave as if they do not interact with each other. At most temperatures (T_1 and T_2), z decreases as the pressure is raised, becoming very small during the condensation process. However, because of the relative incompressibility of the liquid, z becomes large at the highest pressures. Mathematical representations of the variation of z with T, P, or ρ and composition are called equations of state, a famous form being that due to J. D. van der Waals. Many equations of state of much higher accuracy have been developed, though the most complete expressions are sufficiently complex that a computer is necessary to actually calculate the properties they can give. *See* GAS; GAS CONSTANT; THERMODYNAMIC PROCESSES; VAN DER WAALS EQUATION; VAPOR PRESSURE.

Mixtures. Mixtures of fluids show the same general density and multiphase behavior as pure fluids, but the composition is an extra variable to be considered. For example, the density differences between the vapor and the liquid phases cause them to have different relative amounts of the components. This difference in composition is the basis of the separation process of distillation, where the vapor will be richer in some components while the liquid will be richer in others. In the distillation refining of petroleum, the vapor is enriched with the components

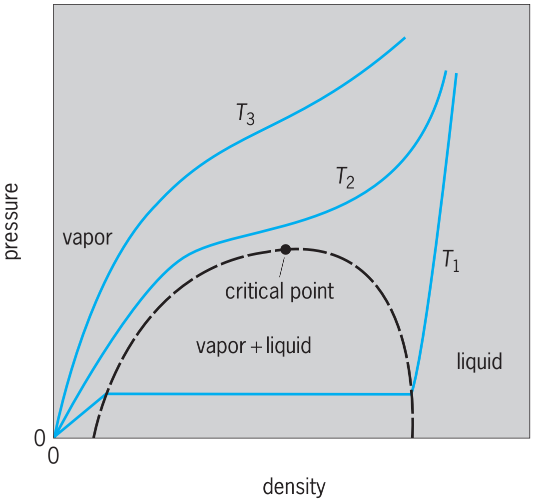


Fig. 3. Behavior of pure-component pressures with density at various temperatures.

of gasoline, while the liquid is richer in heavier oils. It is also possible for mixtures of liquids to be partially or nearly wholly immiscible, as are water and oil. The separation process of liquid extraction, used in some metal-purification systems and chemical-pollution-abatement processes, depends on different preferences of chemical solutes for one liquid phase or the other. Both vapor-liquid and liquid-liquid systems can show critical points where the densities and compositions become the same. In the single-phase conditions near the critical point, the great variation of density with small changes in pressure can be used to selectively remove components from sensitive substances such as foods and pharmaceuticals. This process (called supercritical extraction) is used to extract caffeine from coffee with carbon dioxide near its critical point. In some systems, it is possible to observe many liquid phases. A well-known example has nine liquid phases made from mixing nine liquid substances. *See* DISTILLATION; EXTRACTION; PHASE RULE.

Fluid interfaces. The usual observation of the presence of more than one fluid phase is the appearance of the boundary or interface between them. This is seen because the density or composition (or both) changes over a distance of a few molecular diameters, and this variation bends or scatters light in a detectable way. At the interface, the molecules feel different forces than in the bulk phases and thus have special properties. Energy is always required to create interface from bulk, the amount per unit area being called the interfacial tension. Water is a fluid with an extremely high vapor-liquid (or surface) tension; this surface tension allows insects to crawl on ponds and causes sprinkler streams to break up into sprays of droplets. One way to measure the surface tension of a fluid is to measure its capillary rise in small tubes. The height of the fluid rise decreases with the tube diameter and increases with the surface tension. Another effect of interfacial tension is the increase of vapor pressure of a component in a very small droplet. Liquid mercury will not evaporate in an open container with an essentially flat interface. However, if it is spilled into little drops, those too small to be easily seen will have their vapor pressure so enhanced that the mercury will vaporize. (This can be a health hazard in chemical and other laboratories.) As the conditions of temperature and composition are changed toward a critical point, the interfacial tension decreases to zero. *See* INTERFACE OF PHASES; SURFACE TENSION.

In mixtures, the molecules respond differently to the interfacial forces, so the interfacial composition is generally different from that of the bulk. This has also been the basis of a separation process. If the difference of composition is great enough and it varies with time and position because of evaporation of one or more of the components, the interfacial forces can push the fluid into motion, as can be observed on the walls of a glass of brandy (the Marangoni effect). Some substances strongly adsorb at the interface because their chemical structure has one part that prefers to be in one phase, such as water, and another part that prefers the other phase, such as oil or air. Such surfactants or detergents help solubilize dirt into wash water, keep cosmetics and other immiscible mixtures together, and form foams when air and soapy water are whipped together. *See* ADSORPTION; DETERGENT; FOAM; SURFACTANT.

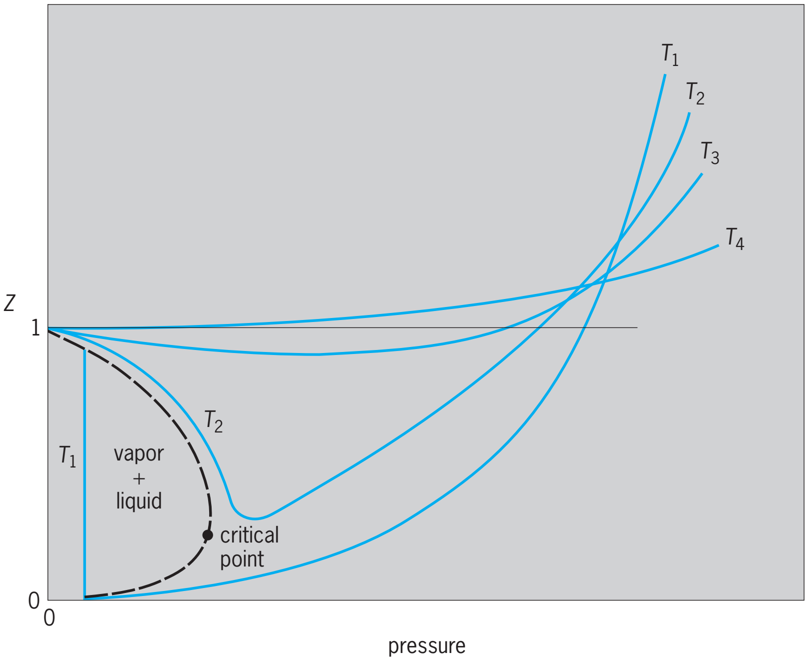


Fig. 4. Behavior of pure-component compressibility factors with pressure at various temperatures.

Transport properties. Besides the relations among pressure, density, temperature, and composition of static or equilibrium fluids, there are also characteristics associated with fluid flow, heat transfer, and material transport. For example, when a liquid or gas flows through a tube, energy must be supplied by a pump, and there is a drop in pressure from the beginning to the end of the tube that matches the rise in pressure in the pump. The pump work and pressure drop depend on the flow rate, the tube size and shape, the density, and a property of the molecules called the viscosity. The effect arises because the fluid molecules at the solid tube wall do not move and there are velocity gradients and shear in the flow. The molecules that collide with one another transfer momentum to the wall and work against one another, in a sort of friction which dissipates mechanical energy into internal energy or heat. The greater the viscosity, the greater the amount of energy dissipated by the collisions and the greater the pressure drop. The viscosity generally increases with the density. In gases, viscosity increases with temperature and decreases with molecular size and attractive forces. In liquids, the effects are opposite; the main effect is due to changes in density. If only chemical constitution and physical state are needed to characterize the viscosity, and if shear stress is directly proportional to velocity gradient, the fluid is called newtonian and the relation for pressure drop is relatively simple. If the molecules are large or the attractive forces are very strong over long ranges, as in polymers, gels, and foods such as bread dough and cornstarch, the resistance to flow can also depend on the rate of flow and even the recent deformations of the substance. These fluids are called non-newtonian, and the relationship of flow resistance to the applied forces can be very complex. *See* NEWTONIAN FLUID; NON-NEWTONIAN FLUID; VISCOSITY.

Another fluid-transport property, thermal conductivity, indicates the ability of a static fluid to pass heat from higher to lower temperature. This characteristic is a function of chemical constitution and physical state in a similar way as is the viscosity. In mixtures, these properties may involve simple or complex dependence on composition, the variation becoming extreme if the unlike species strongly attract each other. The values of both properties increase rapidly near a critical point. *See* CONDUCTION (HEAT); HEAT TRANSFER.

Finally, the ability of molecules to change their relative position in a static fluid is called the diffusivity. This is a particularly important characteristic for separation processes whose efficiency depends on molecular motion from one phase to another through a relatively static interface, or on the ability of some molecules to move faster than others in a static fluid under an applied force. The diffusivity is highest for the smallest molecules with the weakest attractive intermolecular forces, decreasing as the size and forces increase. The diffusivity in gases increases with temperature and decreases with density; often, the product of the diffusivity and the density divided by the product of the viscosity and the absolute temperature is nearly constant. In liquids the diffusivity tends to be inversely proportional to viscosity of solvent. *See* DIFFUSION; GAS; LIQUID.

John P. O'Connell

Bibliography. P. W. Atkins, *Physical Chemistry*, 6th ed., 1997; G. M. Barrow, *Physical Chemistry*, 6th ed., 1996; I. N. Levine, *Physical Chemistry*, 5th ed., 2000; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; J. M. Prausnitz et. al., *Molecular Dynamics of Fluid-Phase Equilibria*, 2d ed., 1985; R. C. Reid, J. M. Prannitz, and B. Poling, *Properties of Gases and Liquids*, 4th ed., 1987.

Fluorescence

Fluorescence is generally defined as a luminescence emission that is caused by the flow of some form of energy into the emitting body, this emission ceasing abruptly when the exciting energy is shut off.

Definition. In attempts to make this definition more meaningful it is often stated, somewhat arbitrarily, that the decay time, or afterglow, of the emission must be of the order of the natural lifetime for allowed radiative transitions in an atom or a molecule, which is about 10^{-8} s for transitions involving visible light. Perhaps a better distinction between fluorescence and its counterpart, a longer-lasting afterglow called phosphorescence, rests not on the magnitude of the decay time per se but on the criterion that the fluorescence decay is temperature-independent. If this latter criterion is adopted, the luminescence emission from such materials as the uranyl compounds and rare-earth-activated solids would be called slow fluorescence rather than phosphorescence. The decay of their luminescence takes place in milliseconds to seconds, rather than in 10^{-8} s, showing that the optical transitions are somewhat "forbidden"; but the decay is temperature-independent over a considerable range of temperature, and it follows an exponential decay law, shown in the equation below, that is to be expected for spontaneous transitions of electrons from an excited state of an atom to the ground state when the atom has a transition probability per unit time $1/\tau_1$ (τ_1 is called the natural radiative lifetime or fluorescence lifetime). In this equation, I is the luminescence intensity at a time t, and I_0 is the intensity when $t = 0$, that is, the time at which the exciting energy is removed. *See* SELECTION RULES (PHYSICS).

$$I = I_0 \exp\left(\frac{-t}{\tau_1}\right)$$

In applying this criterion, it is necessary to take into account a restriction that arises because all luminescent systems ultimately lose efficiency or are "quenched" at elevated temperatures, each system having its own characteristic temperature for the onset of this so-called thermal quenching. Quenching sets in because increase of temperature makes available other competing atomic or molecular transitions that can depopulate the excited state and dissipate the excitation energy nonradiatively. Hence, even for a fluorescence, a temperature dependence of decay time will be observed at temperatures where thermal quenching becomes operative, because $1/\tau_{\mathrm{obs}} = 1/\tau_1 + 1/\tau_{\mathrm{diss}}$, where τ_{obs} is the observed fluorescence lifetime and $1/\tau_{\mathrm{diss}}$ is the probability for a competing dissipative transition.

In the literature of organic luminescent materials, the term fluorescence is used exclusively to denote a luminescence which occurs when a molecule makes an allowed optical transition. Luminescence with a longer exponential decay time, corresponding to an optically forbidden transition, is called phosphorescence, and it has a different spectral distribution from the fluorescence. *See* PHOSPHORESCENCE.

Materials. Fluorescence can be excited in gases, liquids, or solids. Solid luminescent materials are commonly called phosphors, regardless of their decay times, which can cover a very wide range, from the order of 5×10^{-9} s for many organic crystalline materials up to 2 s for the europium-activated strontium silicate phosphor.

Applications. Fluorescent materials with decay times between 10^{-9} and 10^{-7} s are used to detect and measure high-energy radiations, such as x-rays and gamma rays, and high-energy particles such as alpha particles, beta particles, and neutrons. These agents produce light flashes (scintillations) in certain crystalline solids, in solutions of many polynuclear aromatic hydrocarbons, or in plastics impregnated with these hydrocarbons. *See* SCINTILLATION COUNTER.

Another application of x-ray-induced fluorescence is in intensifying screens used in medical radiography. Such screens, placed in contact with radiographic film, considerably decrease the x-ray exposure needed to form an image because of the added photographic effect of the fluorescence. Blue-fluorescent calcium tungstate and rare-earth-activated phosphors are normally used.

Electronic displays. The fluorescence excited by electron-beam bombardment of a phosphor coating on a cathode-ray-tube face produces the images seen in television. For this purpose, phosphors with short decay times must be used, because each phosphor picture element is scanned very briefly by the electron beam approximately 30 times per second, and the luminescence from each excitation must decrease to a low level that will not interfere with the next signal. Phosphors with decay times as long as a few hundredths of a second can be tolerated in this application, although shorter decay times are more advantageous. *See* CATHODE-RAY TUBE; PICTURE TUBE.

Vacuum fluorescence display devices also make use of fluorescence excited by an electron beam in a flat structure resembling a triode vacuum tube. In these devices the phosphor is coated on the anodes of the tube, which are shaped to form the segments of various digital or alphanumeric figures, and the excitation of each segment is controlled by the potential applied to that anode. *See* CATHODOLUMINESCENCE; ELECTRONIC DISPLAY.

Visual effects. Laundry brighteners are another application of fluorescence. These materials are colorless organic phenyl-based dyes. They are excited by the small amount of ultraviolet light present in daylight to give a blue fluorescence, thus compensating for the slight absorption of blue light responsible for the cream color of most natural bleached textiles. The fluorescence of other dyes and inorganic phosphors applied to fabrics or other substrates is used to obtain special visual effects for advertising, theatrical, and other display purposes, under both daylight and ultraviolet excitation. *See* DYE.

Research, diagnosis, and prospecting. Fluorescence is extensively used for research and diagnosis in microbiology and medicine, where fluorescent dyes, such as fluorescein, rhodamine, and aniline blue, are used to label and trace proteins. Among other applications of fluorescence are the use of water-soluble fluorescent dyes to trace the flow of hidden currents, and as an aid in prospecting for certain minerals, sceelite (calcium tungstate) giving a blue fluorescence and willemite (zinc silicate) giving a green fluorescence under short-wave ultraviolet excitation. *See* FLUORESCENCE MICROSCOPE; IMMUNOFLUORESCENCE; SCHEELITE; WILLEMITE.

Fluorescent lamps. The so-called fluorescent lamps employ the luminescence of gases and solids in combination to produce visible light. A fluorescent lamp consists of a glass tube filled with a low-pressure mixture of argon gas and mercury vapor, coated on the inside surface with a luminescent powder or blend of such powders, and having an electrode at each end. An electrical discharge is passed through the gas between the two electrodes. Bombardment by the electrons and ions of the discharge excites the mercury atoms to higher energy states, from which they return to lower states by radiative transitions, producing both visible and ultraviolet light emission. At the low pressures employed, approximately half of the electrical energy input to the lamp is converted into the 253.7-nanometer radiation characteristic of the mercury atom.

The phosphor coating is chosen for the efficiency with which it is excited by this wavelength of ultraviolet and for the color of visible luminescence that is desired. Lamp phosphors are predominantly alkaline-earth halophosphates, symbolized by $(Ca,Sr)_5(PO_4)_3(F,Cl)$, activated by antimony and manganese. The phosphors operate on the principle of sensitized luminescence; antimony, the sensitizer, is the primary absorber of the 253.7-nm radiation. Some of the excited antimony atoms fluoresce in the blue-green spectral region, while others transfer their excitation energy to manganese atoms, which fluoresce in the yellow-red region of the spectrum. The net color of the lamp fluorescence is thus some shade of white determined by the exact composition of the phosphor or blend of phosphors employed.

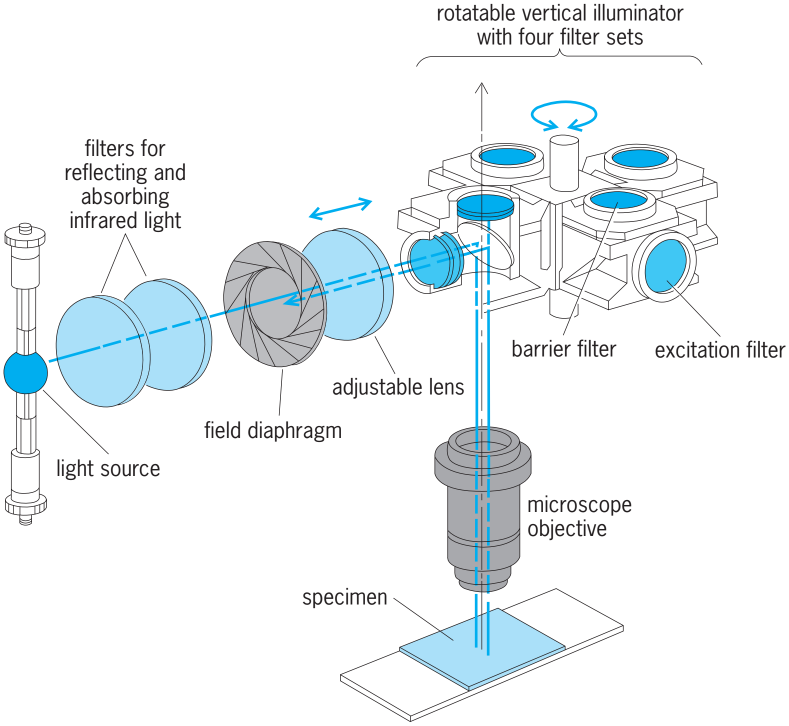
The small amount of visible light generated by the discharge itself is largely transmitted by the phosphor coating and adds slightly to the luminous output of the lamp. Luminous efficiencies of fluorescent lamps are considerably higher than those of incandescent lamps, which unavoidably convert into heat the major portion of the electrical energy supplied to them. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; FLUORESCENT LAMP; LUMINESCENCE.

James H. Schulman

Bibliography. B. DiBartolo (ed.), *Luminescence of Inorganic Solids*, 1978; G. G. Guilbault (ed.), *Practical Fluorescence*, 1990; K. Nassau, *Physics and Chemistry of Color*, 1983.

Fluorescence microscope

An instrument for the observation and study of microscopic specimens that absorb light and emit fluorescence. The ability of substances to emit light when excited by a strong light source is called luminescence, which encompasses the phenomena of phosphorescence and fluorescence. If the secondary emission continues for some time (milliseconds) after the light source has been turned off, it is referred to as phosphorescence. The ability to emit light only very briefly (nanoseconds) after the absorption of light is termed fluorescence. Many specimens obtained from plants, animals, and minerals have the ability to emit a weak to strong fluorescence when excited with appropriate light and without being specifically treated or stained. That phenomenon is called autofluorescence. In most



Schematic diagram of a fluorescence illuminator, the Ploemopak, with four filter sets.

cases, in order to obtain specific and meaningful fluorescence, staining with fluorescing dyes called fluorophores or fluorochromes is necessary. Fluorescence microscopy is a highly sensitive method, since often minute quantities of a fluorophore can be visualized with good microscopic contrast. In appropriate applications, brightly fluorescing images can be observed against a dark background. Individual fluorophores have different absorption and emission spectra and a different quantum efficiency (the ratio between the energy absorbed and the energy emitted), factors that must be considered for optimum fluorescence. *See* BIOLUMINESCENCE; FLUORESCENCE; LUMINESCENCE; PHOSPHORESCENCE.

Microscope components. The basics for fluorescence microscopy are the light source necessary to illuminate the specimen and the optics needed to observe the fluorescence. In addition, filters must be used to single out appropriate excitation and emission wavelengths. Excitation filters select out a limited range of excitation wavelengths from the light source that corresponds to the absorption spectrum of the fluorochrome. Barrier filters separate emitted light from unabsorbed exciting light. Fluorescence can then be observed by eye, photographed, measured by a photomultiplier, or recorded by a television camera. Two types of illumination are used in fluorescence microscopy: transmitted and incident. The earliest fluorescence microscopes relied on transmitted illumination, which generally used a dark-ground condenser to facilitate the separation of fluorescent and exciting light. With the development of epi-illumination, the exciting light reaches the preparation from above by way of a dichroic mirror and the objective, which at the same time acts as a condenser.

Epi-illumination by means of a vertical illuminator, which permits the excitation with more wavelengths, is known as a fluorescence illuminator (see **illus.**), and has become the routine instrument for fluorescence microscopy. Its dichroic has a special interference coating that reflects light shorter than a specified wavelength and transmits that of longer wavelengths. By placing the dichroic mirror at a 45° angle in the light path above the objective, the exciting, or shorter, wavelengths are reflected onto the preparation, and the emitted, or longer, wavelengths are transmitted in the direction of the eyepieces. Unlike microscopes with transmitted illumination, which require focusing of both the substage condenser and the objective, the fluorescence microscope with a fluorescence illuminator requires focusing of the objective only. In addition, a combination of epi-illumination fluorescence microscopy and transmitted microscopy is feasible.

Fluorescence illuminators have been manufactured with four different sets of dichroic mirrors, excitation filters, and barrier filters, each combination appropriate for the excitation of a specific fluorochrome. They are particularly useful in immunofluorescence microscopy, where often two or more fluorochromes are being used in one specimen. Optimum images can be obtained when low-power, high-numerical-aperture objectives are used, because the intensity of fluorescence is directly proportional to the numerical aperture of the objective (in incident illumination, to the fourth power of the numerical aperture) and image brightness is inversely proportional to the square of the total magnification. In general, simpler optics are preferred in fluorescence microscopy because they will have a higher transmission for fluorescent light and will produce less autofluorescence. Only ultraviolet exciting light demands special, quartz optics. *See* OPTICAL MICROSCOPE; OPTICAL PUMPING.

Confocal laser-scanning fluorescence microscopy. In laser-scanning microscopy, the object is not illuminated as a whole but is scanned step by step with a laser-illuminated spot. From each point the fluorescence is measured and, after analog-to-digital conversion, stored as a matrix in computer memory. The main advantage of laser-scanning microscopy over conventional methods in flourescence microscopy is the point-by-point illumination. Most stray light is avoided, and so the level of autofluorescence from optical parts and immersion oil is extremely low. The concentration of the laser light on a very small spot results in a high level of exciting energy per unit of area. That permits the investigation of very weakly fluorescing structures and the use of very low magnification (25 times), which is necessary, for example, in mapping the brain. The amount of fading is very low because of the extremely short exposure times. In the confocal mode, which uses a pinhole in the illumination pathway and a corresponding pinhole in front of the light detector, the fluorescence from above and below the selected focal plane is almost completely eliminated from the image. That removes most of the glare experienced in conventional fluorescence microscopy of thicker specimens. Confocal fluorescence microscopy allows optical image sectioning of the specimen and, after combining the multiple images at different focal levels of a specimen, a computerized reconstruction of three-dimensional images of a section. *See* CONFOCAL MICROSCOPY; LASER.

Applications. The primary application of fluorescence microscopy is in the field of medicine, where diagnostic tests have been developed that use monoclonal antibodies to which a red, green, or blue fluorescent dye has been attached. The dyes thus reveal various components in the specimen, such as bacteria, viruses, or macromolecules like deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Many fluorescent cytochemical staining reactions have been developed to detect other components in plant and animal tissue. Outside the field of biology, fluorescence microscopy can analyze precisely the spectral wavelength range of the emitted autofluorescence of minerals and materials used in integrated electronic circuits, thereby helping identify specific substances or chemical impurities. *See* ANIMAL VIRUS; IMMUNOFLUORESCENCE; MEDICAL BACTERIOLOGY.

Johan S. Ploem

Fluorescent lamp

A lamp that produces light largely by conversion of ultraviolet energy from a low-pressure mercury arc to visible light. Phosphors, chemicals that absorb radiant energy of a given wavelength and reradiate at longer wavelengths, produce most of the light provided by fluorescent lamps. *See* PHOSPHORESCENCE.

The lamp consists of a glass tube containing two electrodes, a coating of activated powdered phosphor, and small amounts of mercury. The glass tube seals the inner parts from the atmosphere. The electrodes provide a source of free electrons to initiate the arc, and are connected to the external circuit through the ends of the lamp. The phosphor converts short-wave ultraviolet energy into visible light. The mercury, when vaporized in the arc, produces the ultraviolet radiation that causes fluorescence. An inert gas, such as argon, krypton, or neon, introduced in small quantities provides the ions that facilitate starting of the lamp (**Fig. 1**). *See* FLUORESCENCE.

Operation. To start the flow of current that forms the arc in a typical fluorescent lamp, free electrons are released into the tube by thermionic emission, field emission, or a combination of both, causing a cathode to emit electrons.

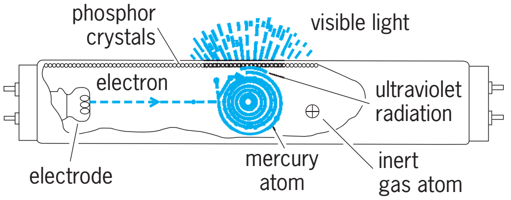


Fig. 1. Parts of a typical fluorescent lamp.

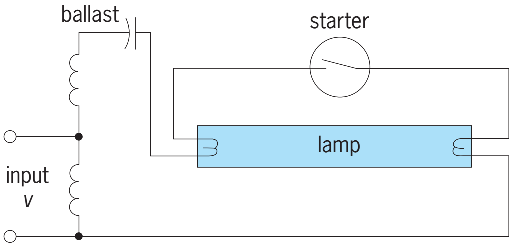


Fig. 2. Preheat or switch-start lamp circuit.

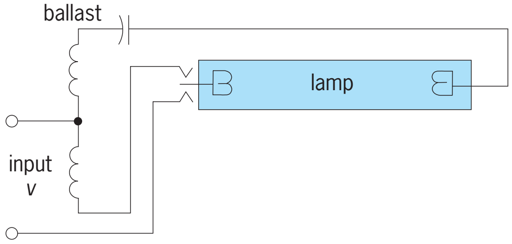


Fig. 3. Instant-start lamp circuit.

In a preheat, rapid-start, or trigger-start type of lamp, current passed through the electrodes causes emission of electrons from the electrodes into the tube. When sufficient electrons are released and a sufficient voltage is applied, the resistance of the gap between the electrodes is low enough to permit striking the arc across the gap. *See* THERMIONIC EMISSION.

The application of a high potential difference across opposite instant-start cathodes in some lamps draws electrons from the negative electrode and attracts them toward the positive electrode. On alternating-current (ac) circuits, the electrodes are alternately negative and positive in each half-cycle, and so both electrodes emit electrons. Instant-start cathodes may be either hot or cold. The hot cathode permits greater lamp current and lower overall lighting costs; it consists of a coiled wire coated with an emissive material that yields electrons freely. The cold cathode is well suited to sign tubing and lamps of special lengths or shapes; it consists of a thimble-shaped iron cup. *See* FIELD EMISSION.

Electrodes designed to heat quickly can be used in conjunction with a moderately high voltage to attract electrons into the tube; this technique is used with rapid-start electrodes. Rapid-start electrodes are continuously heated during lamp operation.

The three types of fluorescent lamps employ either the preheat, instant-start, or rapid-start electric circuits. All include a ballast, whose primary function is to control the amount of current that flows through the lamp. Without a ballast, a fluorescent lamp would burn out instantly because there would be no impedance to limit the current. Another necessary function of the ballast is to provide the proper voltage to start the lamp. For example, a typical 40-W rapid-start ballast must provide a continuous cathode

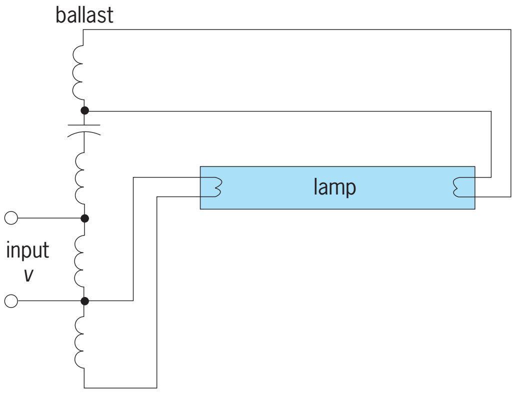


Fig. 4. Rapid-start or trigger-start lamp circuit.

heating voltage of 3.5 V during starting and operation in order to obtain optimum lamp performance. Many circuit variations are used for the operation of one or more lamps from a single ballast, with and without circuit elements for the correction of power factor.

Preheat circuit. In the preheat circuit (**Fig. 2**), a starting switch, usually an automatic starter, is used to heat the electrodes. The most common starter employs a small argon glow tube with one fixed electrode and one electrode made of a bent bimetallic strip. When the circuit is energized, a low-current glow discharge forms in the starter; the heat from this glow is sufficient to expand the bimetallic electrode until it contacts the fixed electrode, forming a short circuit through the starter. The full output voltage of the ballast then causes current to flow through the lamp electrodes, heating them and causing them to emit electrons. The starter cools because the glow is no longer present, and the switch opens, impressing full ballast voltage between the lamp electrodes. If there are enough electrons in the tube, the arc is formed; if not, the glow switch process is repeated. This process may require several repetitions, accounting for the delay and flickering at the starting of preheat lamps. When the arc has formed and full lamp current is flowing, the ballast absorbs about half its initial voltage, and there is not sufficient voltage to cause the glow switch to operate; it then becomes an inactive circuit element. *See* NEON GLOW LAMP.

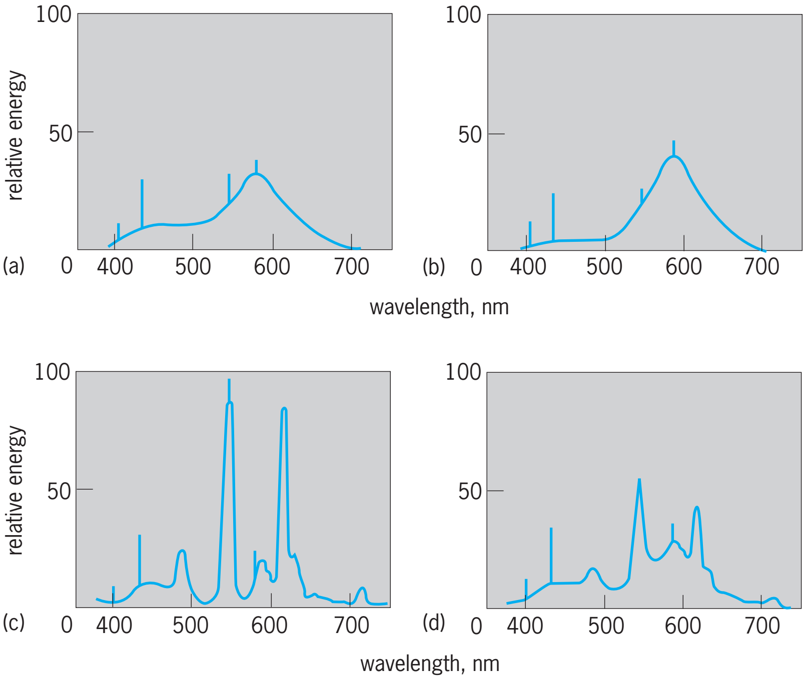


Fig. 5. Spectral energy distribution curves for fluorescent lamps with mercury emission-line spectrum. (*a*) Cool white halophosphate. (*b*) Warm white halophosphate. (*c*) Rare-earth–activated 4100 K (7000○F) triphosphor. (*d*) Triphosphor-halophosphate combination. (*General Electric Co.*)

Preheat fluorescent lamps are commercially available in lengths from 6 in. to 5 ft (15 cm to 1.5 m), with wattages from 4 to 90 W. Some preheat lamps have a self-contained ballast, which makes them useful as a portable light source. Longer lamps or lamps of higher wattage generally use instant-start or rapid-start circuits, which have more favorable starting characteristics. Each size of fluorescent lamp requires a ballast designed to match the lamp's requirements in terms of voltage, current, and wattage.

Instant-start circuit. In the instant-start circuit (**Fig. 3**), the ballast voltage is much higher than in the preheat circuit. Immediately upon energizing the circuit, this voltage attracts enough electrons into the tube to form the arc. For a given lamp wattage and current, the ballast must absorb a higher voltage than with preheat ballasts; hence the ballast is usually larger and dissipates greater wattage. Lampholders for instant-start circuits are connected so that the ballast is disconnected when the lamp is removed, eliminating the hazard of high voltage during maintenance operations.

Instant-start lamps for general lighting purposes are available in lengths from 24 in. to 8 ft (0.6 to 2.4 m), with wattages from about 15 to 75 W. Shorter lamps are not usually economical for general lighting service, because they require larger and more costly ballast than preheat circuits. Higher wattages would require overly large ballasts and cathodes of costly construction.

Rapid-start circuit. The rapid-start ballast, which has become the most popular type of ballast for domestic use in the United States, contains transformer windings that continuously provide the required voltage and current for electrode heating (**Fig. 4**). When the circuit is energized, these windings quickly heat the electrodes, releasing enough electrons into the tube for the lamp to arc from the voltage of the secondary windings. This combination of heat and moderately high voltage permits quick lamp starting with smaller ballasts than those for instant-start lamps, eliminates the annoying flicker associated with the starting of preheat lamps, and eliminates the starter and its socket from the lighting system.

The continuously heated cathode of the rapid-start lamp is better adapted to higher lamp currents and wattages. Therefore, rapid-start lamps are available in wattages up to 215 W for 8-ft (2.4-m) lamps. Compact twin tubes have been developed that are compatible with rapid-start operation. They use smaller diameters (0.4–0.6 in. or 10–15 mm), range in wattage from 5 to 40 W, and vary in length from 6 to 22.5 in. (15 to 57 cm).

Special ballast and circuits are available that permit the economical dimming and flashing of rapid-start lamps, providing a range of applications and control that was not possible with previous fluorescent lamp types. Rapid-start lamps are commonly used in flashing signs, and in residential and commercial lighting where continuously variable illumination levels are desired.

High-frequency fluorescent lighting. Fluorescent lamps are usually operated on ac circuits with a frequency of 60 Hz. However, higher frequencies permit higher-efficacy lamp operation along with ballasts of lower power dissipation per watt. Consequently, systems have been developed for the operation of fluorescent lamps at frequencies from 360 to 50,000 Hz. The most important high-frequency ballasts operate the lamps in the 25-kHz range, are lighter in weight, can have flicker-free light output, and will continue to be more cost-effective as the cost of electricity increases. Electronic technology permits these ballasts to be made with more features, such as dimming and low-cost remote controls, enhancing their overall performance value.

Lamp colors. Fluorescent lamps provide light at several times the efficacy of incandescent lamps, the exact ratio depending on the fluorescent lamp color. Lamp color is determined by the selection of chemicals used in the phosphors; various chemicals respond to the ultraviolet energy in the arc by producing different colors of light. Several types of essentially white fluorescent lamps are available commercially, as well as a range of tinted and saturated colors.

One of the earliest fluorescent whites was the so-called daylight lamp, which produced a bluish-white light of poor color rendition. The most widely used fluorescent color is called cool white; its lighted appearance is whiter than that of the daylight lamp and produces more light per watt. However, rare-earth–activated phosphors are also efficient and have substantially higher color rendition properties. **Figure 5** illustrates some typical spectral energy distributions.

Lamp shapes. The conventional fluorescent lamp is an elongated round tube. Lamp lengths vary from 6 to 96 in. (15 cm to 2.4 m). The highest-wattage lamps have indented grooves along the tube length. Compact fluorescent lamps come in circular, U-shaped, rectangular, twin-tube, quad, and other configurations.

Performance. Fluorescent lamps, commercially available since 1940, have been constantly improving in performance. Phosphor improvements have increased lumen outputs, and electrode designs now yield lamp life values of 20,000 h or more. An important measure of lamp performance is luminous efficacy, the quotient of the total luminous flux divided by the total lamp power input. The F40 system had a luminous efficacy of about 50 lumens per watt in 1950, and has reached values of 90 lumens per watt. The longer 96-in. (2.4-m) lamps, when used with the latest-technology ballasts, are expected to achieve efficacy values of about 100 lu/W. For proper low-temperature performance (−20°F or −29°C), fluorescent lamps are enclosed within jackets to assure high light output. *See* VAPOR LAMP.

Edward E. Hammer

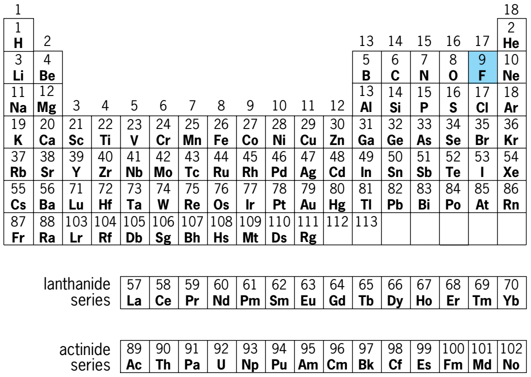
Bibliography. E. E. Hammer, High frequency characteristics of fluorescent lamps up to 500 kHz, *J. Illum. Eng. Soc.*, 16(1):52–61, 1987; E. E. Hammer and T. K. McGowan, Characteristics of various F40 systems at 60 Hz and high frequency, *IEEE Trans. Ind. Appl.*, IA-21:11–16, 1985; *Industry Applications Society Conference Record*, San Francisco, IEEE 82- CH-1817-6, 1982; C. Meyer and H. Nienhuis, *Discharge Lamps*, 1989.

Fluorine

A chemical element, F, atomic number 9, the member of the halogen family that has the lowest atomic number and atomic weight. Although only the isotope with atomic weight 19 is stable, the other, radioactive isotopes between atomic weight 17 and 22 have been artificially prepared. Fluorine is the most electronegative element, and by a substantial margin the most chemically energetic of the nonmetallic elements. *See* PERIODIC TABLE.

Properties. The element fluorine is a pale yellow gas at ordinary temperatures. The odor of the element is somewhat in doubt. Some physical properties are listed in the table. The reactivity of the element is so great that it will react readily at ordinary temperatures with many other elementary substances, such as sulfur, iodine, phosphorus, bromine, and most metals. Since the products of the reactions with the nonmetals are in the liquid or gaseous state, the reactions continue to the complete consumption of the fluorine, frequently with the evolution of considerable heat and light. Reactions with the metals usually form a protective metallic fluoride which blocks further reaction, unless the temperature is raised. Aluminum, nickel, magnesium, and copper form such protective fluoride coatings.

Fluorine reacts with considerable violence with most hydrogen-containing compounds, such as water, ammonia, and all organic chemical substances



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

whether liquids, solids, or gases. The reaction of fluorine with water is very complex, yielding mainly hydrogen fluoride and oxygen with less amounts of hydrogen peroxide, oxygen difluoride, and ozone. Fluorine displaces other nonmetallic elements from their compounds, even those nearest fluorine in chemical activity. It displaces chlorine from sodium chloride, and oxygen from silica, glass, and some ceramic materials. In the absence of hydrofluoric acid, however, fluorine does not significantly etch quartz or glass even after several hours at temperatures as high as 390°F (200°C).

Fluorine is a very toxic and reactive element. Many of its compounds, especially inorganic, are also toxic and can cause severe and deep burns. Care must be taken to prevent liquids or vapors from coming in contact with the skin or eyes.

Natural occurrence. At an estimated 0.065% of the Earth's crust, fluorine is roughly as plentiful as carbon, nitrogen, or chlorine, and much more plentiful than copper or lead, though much less abundant than iron, aluminum, or magnesium. Compounds whose molecules contain atoms of fluorine are widely distributed in nature. Many minerals contain small amounts of the element, and it is found in both sedimentary and igneous rocks.

Uses. Fluorine-containing compounds are used to increase the fluidity of melts and slags in the glass and ceramic industries. Fluorspar (calcium fluoride) is introduced into the blast furnace to reduce the viscosity of the slag in the metallurgy of iron. Cryolite, Na_2AlF_6, is used to form the electrolyte in the metallurgy of aluminum. Aluminum oxide is dissolved in this electrolyte, and the metal is reduced electrically from the melt. The use of halocarbons containing fluorine as refrigerants was patented in 1930, and these volatile and stable compounds found a market in aerosol propellants as well as in refrigeration and air-conditioning systems. However, use of fluorocarbons as propellants has declined sharply because of concern over their possible damage to the ozone layer of the atmosphere. A use for fluorine that became prominent during World War II is in the enrichment of the fissionable isotope ^{235}U; the most important process employed uranium hexafluoride. This stable, volatile compound was by far the most suitable material for isotope separation by gaseous diffusion.

While consumers are mostly unaware of the fluorine compounds used in industry, some compounds have become familiar to the general public through minor but important uses, such as additives to toothpaste and nonsticking fluoropolymer surfaces on frying pans and razor blades (for example, Teflon).

Physical properties of fluorine

Property	Value
Atomic weight	18.998403
Boiling point, °C	−188.13
Freezing point, °C	−219.61
Critical temperature, °C	−129.2
Critical pressure, atm*	55
Density of liquid at b.p., g/ml	1.505
Density of gas at 0°C + 1 atm*, g/liter	1.696
Dissociation energy, kcal/mol	36.8
Heat of vaporization, cal/mol	1510
Heat of fusion, cal/mol	121.98
Transition temperature (solid), °C	−227.61

*1 atm = 101.325 kilopascals.

Compounds. In all fluorine compounds the high electronegativity of this element suggests that the fluorine atom has an excess of negative charge. It is convenient, however, to divide the inorganic binary fluorides into saltlike (ionic lattice) nonvolatile metallic fluorides and volatile fluorides, mostly of the nonmetals. Some metal hexafluorides and the noble-gas fluorides show volatility that is frequently associated with a molecular compound. Volatility is often associated with a high oxidation number for the positive element.

The metals characteristically form nonvolatile ionic fluorides where electron transfer is substantial and the crystal lattice is determined by ionic size and the predictable electrostatic interactions. When the coordination number and valence are the same, for example, BF_3, SiF_4, and WF_6, the binding between metal and fluoride is not unusual, but the resulting compounds are very volatile, and the solids show molecular lattices rather than ionic lattice structures. For higher oxidation numbers, simple ionic lattices are less common and, while the bond between the central atom and fluorine usually still involves transfer of some charge to the fluorine, molecular structures are identifiable in the condensed phases.

In addition to the binary fluorides, a very large number of complex fluorides have been isolated, often with a fluoroanion containing a central atom of high oxidation number. The binary saltlike fluorides show a great tendency to combine with other binary fluorides to form a large number of complex or double salts.

The fluorine-containing compounds of carbon can be divided into fluorine-containing hydrocarbons and hydrocarbon derivatives (organic fluorine compounds) and the fluorocarbons and their derivatives. The fluorine atom attached to the aromatic ring, as in fluorobenzene, is quite unreactive. In addition, it reduces the reactivity of the molecule as a whole. Dyes, for example, that contain fluorine attached to the aromatic ring are more resistant to oxidation and are more light-fast than dyes that do not contain fluorine. Most aliphatic compounds, such as the alkyl fluorides, are unstable and lose hydrogen fluoride readily. These compounds are difficult to make and to keep and are not likely to become very important. *See* FLUOROCARBON; HALOGEN ELEMENTS. Irving Sheft

Organic compounds. The carbon compounds containing fluorine belong to several classes, depending on what other substituents besides fluorine are present. The physical properties and chemical reactivity of organic molecules containing fluorine are quite different when compared to the same molecules containing other halogen atoms, such as chlorine. This is due, in part, to a unique combination of the properties of fluorine, which include its small atomic size and high electronegativity. Stepwise replacement of several or all of the hydrogen atoms or other substituents attached to carbon is possible.

Many methods are available for creating a carbon-to-fluorine bond. A widely used method is to exchange a chlorine attached to carbon by reacting the compound with hydrofluoric acid. Elemental fluorine, which is very highly reactive, has also been used to prepare fluorine-containing compounds from a wide variety of organic compounds. The unusual property imparted to an organic molecule by fluorine substitution has led to the development of compounds that fulfill specific needs in refrigeration, medicine, agriculture, plastics, textiles, and other areas.

Fluoroolefins. These are a class of unsaturated carbon compounds containing fluorine; that is, they have a $C{=}C$ in addition to other substituents. A typical fluoroolefin is tetrafluoroethylene ($F_2C{=}CF_2$). It is prepared from chlorodifluoromethane ($CHClF_2$), which loses HCl upon heating to produce $F_2C{=}CF_2$.

Many fluoroolefins combine with themselves or other olefins by the process of polymerization. Thus, polymerization of $F_2C{=}CF_2$ yields the polymer polytetrafluoroethylene (PTFE). This remarkable solid substance has outstanding physical and chemical properties. Nonstick polytetrafluoroethylene surfaces are used in kitchen utensils, bearings, skis, and many other applications. Since polytetrafluoroethylene is very viscous above its melting point, special methods have to be used for fabrication. For this reason, copolymers of tetrafluoroethylene with such olefins as ethylene have been developed. The chemical resistance of these copolymers is less than that of perfluorinated polymers. To obtain polymers with desired properties, the chemical processes to make them are carried out under rigorously controlled conditions. *See* COPOLYMER; POLYFLUOROOLEFIN RESINS; POLYMER; POLYMERIZATION.

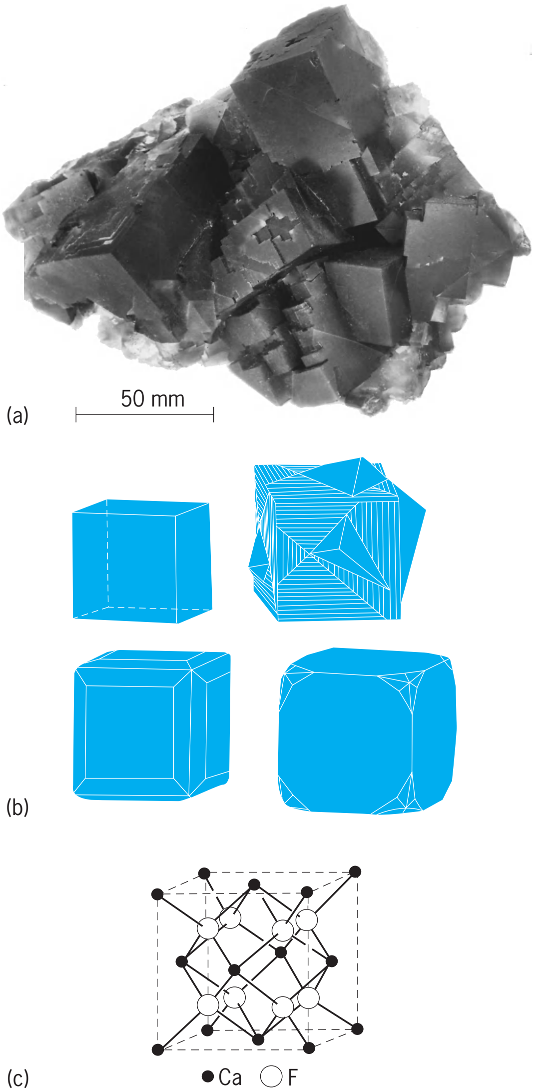
There are many oxygen-containing fluorocarbons such as ethers, acids, ketones, and alcohols. Simple, fluorinated ethers are compounds of the type R-O-R, where R is a fluorinated alkyl group. The simple compound perfluoro ether (F_3COCF_3) is an analog of dimethyl ether. *See* ETHER.

Organofluorine chemicals offer some unique properties and solutions. In addition to the applications mentioned above, they are used in dyes, surfactants, pesticides, blood substitutes, textile chemicals, and biologically active compounds. *See* FLUOROCARBON; HALOGENATED HYDROCARBON. V. N. M. Rao

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Fluorite

A mineral of composition CaF_2 and an important structure type with space group Fm3m, cell edge = 0.546 nm. The fluorite structure type is written RX_2, where the *R* cations are in cubic (eightfold) coordination by the *X* anions. Alternatively, the *X* atoms are tetrahedrally (fourfold) coordinated by the *R* atoms. The fluorite arrangement may occur when the ratio of $r(R)$, the ionic radius of the *R* atoms, to $r(X)$, the ionic radius of the *X* atoms, is equal to or greater than 0.73. The arrangement XR_2 is termed the antifluorite arrangement. Here the anions are eight-coordinated and the cations are four-coordinated. Halides of larger divalent cations, oxides of larger quadrivalent (4+) cations, some intermetallic compounds, and XR_2 alkali univalent oxides frequently belong to the fluorite structure type. Approximately 100 synthetic compounds are known to have this arrangement; distortions may result in similar arrangements but with lower symmetry, such as in ZrO_2, which is monoclinic. A vast number of complex structures can be derived from the fluorite arrangement by ordering vacancies and different cations or anions over the fluorite atomic positions.



Fluorite. (*a*) Crystals from Elizabethtown, Illinois (*American Museum of Natural History specimens*). (*b*) Crystal habits (*after C. S. Hurlbut, Jr., and C. Klein, Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*). (*c*) Structure (*after C. S. Hurlbut, Jr., and C. Klein, Manual of Mineralogy, 21st ed., Wiley, 1993*).

The most abundant fluorine-bearing mineral, fluorite, occurs as cubes or compact masses and more rarely as octahedra with complex modifications (see **illus.**). The cube surfaces frequently show lineage features. Fluorite has a perfect octahedral cleavage, hardness 4 (Mohs scale), and specific gravity 3.18. The color is extremely variable, the most common being green and purple; but fluorite may also be colorless, white, yellow, blue, or brown. Colors may result from the presence of impurity ions such as rare earths or Mn^{4+}, hydrocarbon inclusions, or lattice defects. Fluorite frequently emits a blue-to-green fluorescence under ultraviolet radiation, especially if rare-earth or hydrocarbon material is present. Some fluorites are thermoluminescent; that is, they emit light when heated.

Fluorite is usually very pure, but some varieties, especially yttrofluorite, are known to carry up to 10 wt % CeF_3, the extra fluorine atoms occurring interstitially in the structure. Fluorite can be recrystallized from crushed material and used as optical glass; its more important uses are as flux material in smelting and as an ore of fluorine and hydrofluoric acid.

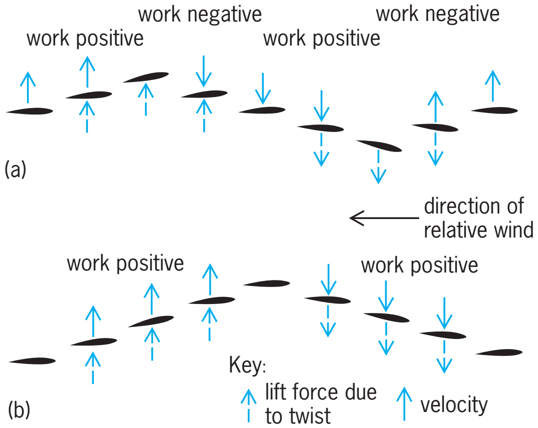
Fluorite occurs as a typical hydrothermal vein mineral with quartz, barite, calcite, sphalerite, and galena. Crystals of great beauty from Cumberland, England, and Rosiclare, Illinois, are highly prized by mineral fanciers. It also occurs as a metasomatic replacement mineral in limestones and marbles. Fluorite is a minor accessory mineral in many other rocks, including granites, granite pegmatites, and nepheline syenites. *See* FLUORINE. Paul B. Moore

Fluorocarbon

Any of the organic compounds in which all of the hydrogen atoms attached to a carbon atom have been replaced by fluorine; also referred to as a perfluorocarbon. Fluorocarbons are usually gases or liquids at room temperature, depending on the number of carbon atoms in the molecule. A major use of gaseous fluorocarbons is in radiation-induced etching processes for the microelectronics industry; the most common one is tetrafluoromethane. Liquid fluorocarbons possess a unique combination of properties that has led to their use as inert fluids for cooling of electronic devices and soldering. Solubility of gases in fluorocarbons has also been used to advantage. For example, they have been used in biological cultures requiring oxygen, and as liquid barrier filters for purifying air. *See* FLUORINE; HALOGENATED HYDROCARBON. V. N. M. Rao

Flutter (aeronautics)

An aeroelastic self-excited vibration with a sustained or divergent amplitude, which occurs when a structure is placed in a flow of sufficiently high velocity. Flutter is an instability that can be extremely violent. At low speeds, in the presence of an airstream, the vibration modes of an aircraft are stable; that is, if the aircraft is disturbed, the ensuing motion will be damped. At higher speeds, the effect of the airstream is to couple two or more vibration modes such that the vibrating structure will extract energy from the airstream. The coupled vibration modes will remain stable as long as the extracted energy is dissipated by the internal damping or friction of the structure. However, a critical speed is reached when the extracted energy equals the amount of energy that the structure is capable of dissipating, and a neutrally stable vibration will persist. This is called the flutter speed. At a higher speed, the vibration amplitude will diverge, and a structural failure will result. *See* AEROELASTICITY.



Wing displacements during a complete cycle. (*a*) Flexural and torsional displacements in phase, resulting in stable motion. (*b*) Torsional displacement lagging 90° behind flexural displacement (torsional displacement in phase with upward flexural velocity), resulting in flutter. (*After W. J. Duncan, Flutter and stability, J. Roy. Aeronaut. Soc., 53:529–549, 1949*)

The **illustration** shows examples of the displacements of a wing cross section during a complete cycle of motion. These diagrams present the wing tip in end view at intervals of one-eighth of a period, and the lift force and work done during the cycle. Part *a* illustrates the wing responding in a stable manner to a disturbance, such as a gust, below the flutter speed, and part *b* the motion during flutter. The illustration shows the physical mechanism of flutter but oversimplifies the phase relationships.

Since the first documented encounter with aerodynamic flutter of an Handley-Page bomber in 1916, the history of flutter has been determined largely by progress in unsteady aerodynamic theory, which is needed to predict both the magnitudes and phases of the aerodynamic forces acting in each of the characteristic aircraft motions. However, sufficient progress was not made until the 1930s. At that time, reliance was on flight testing for solutions to flutter problems. In 1938 a carefully planned flight flutter test of a four-engine Junkers JU 90 aircraft ended in a tragic accident in which the pilot and all the engineers aboard were killed. It became apparent that the many potential dangers associated with flutter must be anticipated before such flight tests. Since then, greater emphasis has been placed on theoretical analysis and wind tunnel tests of dynamically scaled flutter models.

Today's aircraft manufacturers have engineering departments whose primary responsibility is flutter safety. Modern flutter analyses involve extensive computations, requiring the use of large-capacity, high-speed digital computers. Flutter engineers contribute to the design by recommending stiffness levels for the structural components and control surface actuation systems and weight distributions on the lifting surfaces, so that the aircraft vibration characteristics will not lead to flutter within the design speeds and altitudes. Examples of aircraft configuration characteristics that may be determined by flutter considerations include engine placement and fuel location and sequencing. Flutter engineers also conduct ground resonance tests to verify calculated vibration characteristics, and then monitor flight flutter tests to demonstrate adequate margins of safety for certification. Finally, these engineers specify maintenance requirements so that a satisfactory design will not deteriorate in service. Examples include rebalancing control surfaces after painting, and putting limits on hinge looseness, that is, free play, in control surfaces.

Much research involves unsteady aerodynamic theory in order to improve prediction methods. The early solutions were obtained for airfoils in two dimensions. Three-dimensional solutions, which began to appear in the 1950s, were made feasible by developments in digital computers. Accurate prediction methods are now available at subsonic and hypersonic speeds. Supersonic methods are being refined and transonic methods are being developed. Perhaps more significant is research involving structural optimization and flutter suppression. Development of filament composite materials, such as graphite/epoxy, permits tailoring of a structure to minimize its weight while providing adequate strength and stiffness to prevent flutter. Control system technology is well established for autopilot design, and has been extended to develop active and adaptive systems that will provide control surface inputs to stabilize an aircraft above the basic aircraft flutter speed. The feasibility of such automatic control systems has been demonstrated in wind tunnel tests. *See* AIRFOIL; AIRFRAME; AUTOPILOT; COMPOSITE MATERIAL; WING; WING STRUCTURE.

William P. Rodden

Bibliography. R. L. Bisplinghoff and H. Ashley, *Principles of Aeroelasticity*, 2d ed., 1962, reprint 1975; E. H. Dowell (ed.), *A Modern Course in Aeroelasticity*, 1995.

Fluvial erosion landforms

Landforms that result from erosion by water flowing on land surfaces. This water may concentrate in channels as streams and rivers or flow in thin sheets and rills down slopes. Essentially all land surfaces are subjected to modification by running water, and it is among the most important surface processes. Valleys are cut, areas become dissected, and sediment is moved from land areas to ocean basins. With increasing dissection and lowering of the landscape, the land area may pass through a series of stages known as the fluvial erosion cycle.

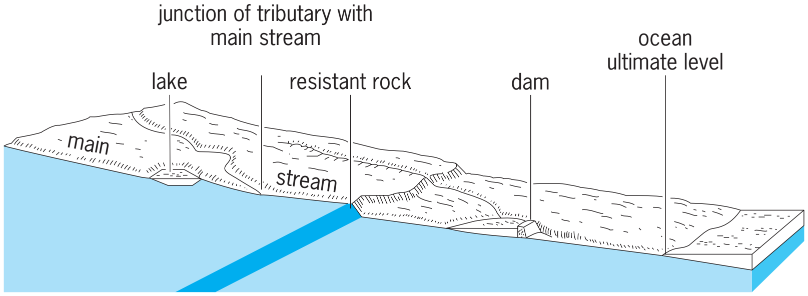


Fig. 1. Differing base levels for streams. (*After L. D. Leet and S. Judson, Physical Geology, 2d ed., Prentice-Hall, 1958*)

The lower limit of stream erosion is called base level. Utimate base level is related to sea level, and can be visualized as an imaginary surface gently inclined upward from sea level toward the interior of a land area. Local or temporary base levels exist where a resistant rock unit crosses a river, where a lake or reservoir occurs along the drainage course, and where a tributary enters a main stream (**Fig. 1**).

Fluvial system. Erosion is an integral part of the fluvial system and, along with transportation and deposition, provides the means by which rivers shape the landscape. The system is complex, and stream activity varies temporally at any one location and spatially up and down the stream course. The main characteristics of a river at a location are channel width, depth, shape and pattern, bed roughness and slope, grain size of load, and flow velocity. These are interrelated closely and are a result of the discharge and the quantity and type of load provided by the drainage basin, and the relative location of the river with respect to base level. Most rivers approach a quasiequilibrium such that the river can adjust its morphology through erosion or deposition as a result of changes in one of the controlling factors or characteristics. Such rivers are termed graded, and are both relatively stable and adjustable. During times of major adjustment the river may be either actively downcutting or aggrading. *See* RIVER; STREAM TRANSPORT AND DEPOSITION.

Fluvial erosion. Flowing water erodes material in the following ways.

1. *Direct lifting*. Loose particles on the stream bottom can be entrained and transported downstream if the stream achieves an appropriate intensity of flow. Main controlling factors are particle size, shape and density, fluid characteristics, and channel depth and slope or flow velocity.

2. *Abrasion and impact*. The solid particles entrained and transported by the stream wear down bedrock portions of the channel through impact with the bed. The bedload of the stream becomes smaller and rounder through this grinding process.

3. *Solution*. In areas of soluble rock the flowing water may dissolve portions of the channel bed. This is a relatively minor process. The enormous amounts of rock material that are carried in solution are mostly contributed to the stream by ground water.

4. *Cavitation*. This refers to the sudden collapse of vapor bubbles which may act locally as a type of hammerblow against the channel bed and be an effective erosive force. Cavitation is restricted to conditions of very great velocity and occurs mainly in waterfalls and rapids or in artificial conduits.

5. *Undercutting*. A stream that erodes laterally by one or more of the above processes may undercut the channel bank or valley wall to the extent that gravity causes the material to slump into the river, where it is moved downstream.

Once in a stream, material may be transported in solution, carried in suspension by turbulent water, or rolled, slid, or bounced along the stream bottom. It may be transported to a basin of sedimentation or, as flow conditions change, may be deposited in the channel or floodplain. *See* DELTA; FLOODPLAIN; SEDIMENTOLOGY.

Features of fluvial erosion. These may range in size from small features of the channel to large regional landscapes.

Channel and valley features. The most distinctive landform is the stream valley, which owes its origin to stream erosion. Valleys range greatly in size and shape, as do the streams that flow in them. They enlarge both through down and lateral cutting by the stream and mass wasting processes acting on the valley sides.

1. *Scour features*. The flow of water and sediment over a bedrock surface commonly leaves it smooth and partially polished as a result of abrasion and scour. The surface may contain potholes, steep-sided depressions scoured out in the bed, as well as other irregularities reflecting the current action of water or the presence of zones of weakness in the rock.

2. *Waterfalls and rapids*. Waterfalls occur where there is a sudden drop in the stream bed. This is often the case where a resistant rock unit crosses the channel and the stream is not able to erode through it at the same rate as the adjacent less resistant rock. Waterfalls also occur where a main valley has eroded down at a faster rate than its tributary valleys which are left hanging above the main stream. With time, waterfalls migrate upstream and are reduced to rapids, and eventually the steepened stream gradient is completely lost.

3. *Cut banks*. Many streams flow in a sinuous or meandering channel, and stream velocity is greatest around the outside of meander bends. Erosion is concentrated in this area, and a steep, cut bank forms. If the river meander impinges against a valley wall, the valley will be widened actively.

4. *Stream terraces*. A stream terrace (**Fig. 2**) represents a former floodplain which has been abandoned as a result of rejuvenation or downcutting by the stream. It is a relatively flat surface with a scarp slope that separates it from the current floodplain or from a lower terrace. Terraces are common features in valleys and are the result of significant changes in the stream system through time.

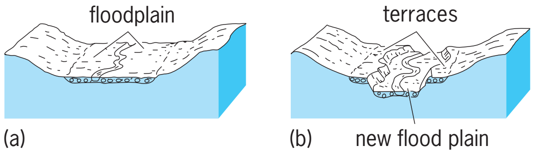


Fig. 2. Development of stream terrace. (*a*) Original floodplain. (*b*) Terrace marks level of abandoned floodplain above level of new floodplain.

Landscape features. Streams and their valleys form a drainage network which reflects the original topography and geologic conditions in the drainage basin.

1. *Drainage patterns*. A dendritic pattern, like that of a branching tree, is the most common and reflects little or no control by underlying earth materials. It develops in areas of sediment cover and flat-lying sedimentary and massive igneous rocks. Where the underlying earth materials are not uniform in resistance, streams develop in the least resistant areas, and the drainage pattern reflects the geology. If the rocks contain a rectangular joint pattern, a rectangular drainage pattern develops; if the rock units are tilted or folded, a trellis pattern of drainage is common. Topography also controls drainage development; parallel and subparallel patterns are common on steep slopes, and a radial pattern develops when streams radiate from a central high area.

2. *Slope morphology and landforms*. Slope morphology is the result of interactions among the surface processes of weathering, mass wasting and flowing water, earth materials beneath the slope, and the degree of erosional activity at the foot of the slope. Slopes in which chemical weathering and surface processes dominate are often characterized by a smooth profile with a convex upper slope and a concave lower slope with or without a straight slope element between them (**Fig. 3***a*). This type of profile is common in humid climates. If chemical weathering is relatively ineffective or erosion is active at the base of the slope, the slope will tend to be steeper and complex, and may contain a near-vertical cliff or free face (Fig. 3*b*). The latter slope element usually reflects a resistant rock unit, and is particularly common in drier areas of the world. *See* WEATHERING PROCESSES.

In areas of variable sedimentary rocks, several characteristic landforms occur depending upon the inclination or dip of the rock units. If the rocks are flat-lying, Fig. 3*b* mesas and buttes are common

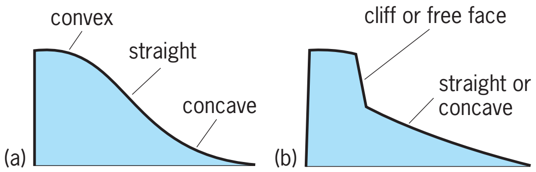


Fig. 3. Two typical slope profiles: (*a*) smooth and (*b*) steep.

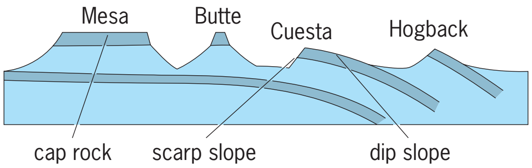


Fig. 4. Landforms developed on flat-lying and dipping sedimentary rocks. Shaded areas represent resistant rock units.

(**Fig. 4**). With a relatively small dip, an asymmetrical landform called a cuesta forms. As the dip becomes larger, the landform becomes more symmetrical and a hogback develops. These landforms may continue for long distances, forming regional escarpments. If they are dissected in a close and regular fashion by V-shaped valleys, the ridge has a triangular shape and is known as a flatiron. *See* EROSION; PLAINS.

W. Hilton Johnson

Bibliography. S. A. Schumm (ed.), *Drainage Basin Morphology*, 1977; S. A. Schumm (ed.), *River Morphology*, 1972; S. A. Schumm, M. P. Mosley, and W. E. Weaver, *Experimental Fluvial Geomorphology*, 1987.

Fluvial sediments

Deposits formed by rivers. An alluvial river is one which flows within its own fluvial sediments, as distinct from one that has incised into the underlying bedrock. A river accumulates deposits because its capacity to carry sediment has been exceeded, and some of the sediment load is deposited. Rivers tend toward a state of dynamic equilibrium, in which they adjust their slope in response to changes in discharge and sediment load. The result is a channel profile that is steep in its source areas but flattens out downstream, and is graded to a slope of zero where the river discharges into a lake or the sea. Fluvial sedimentary accumulations range from temporary bars deposited on the insides of meander bends as a result of a loss of transport energy within a local eddy, to deposits tens to hundreds of meters thick formed within major valleys or on coastal plains as a result of the response of rivers to a long-term rise in base level or to the uplift of sediment source areas relative to the alluvial plain. Both these processes perturb the graded profile so that it tends to rise, creating space, or accommodation, for sediment. The same processes control the style of rivers and the range of deposits that are formed, so that a study of the deposits may enable the geologist to reconstruct the changes in controlling factors during the accumulation of the deposits. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; RIVER; STREAM TRANSPORT AND DEPOSITION.

Coarse debris generated by mechanical weathering, including boulders, pebbles, and sand, is rolled or bounced along the river bed and is called bedload. The larger particles may be moved only infrequently during major floods. Finer material, of silt and clay grade, is transported as a suspended load, and there may also be a dissolved load generated by chemical weathering. Mass movement of large volumes of sediment by sediment gravity flows, typically debris flows, may occur when rare flash floods mobilize debris that may have been accumulating in source areas for some time. Whereas the volume of sediment tends to increase downstream within a drainage system, as tributaries run together, the grain size generally decreases as a result of abrasion and selective transport. This downstream grain-size decrease may assist in the reconstruction of transport directions in ancient deposits where other evidence of paleogeography has been obscured by erosion or tectonic change. *See* MASS WASTING.

Types of river. River type may be described by two main variables, sinuosity and channel multiplicity. These variables combine to form four end-member styles, discussed below, although there are many examples of rivers showing various styles intermediate between these end members.

Braided rivers. These typically occur in areas of high sediment load and variable discharge. They consist of several or many branching, unstable channels of low sinuosity, and are characterized by abundant coarse bedload, forming bars, islands, and channel-floor deposits. The channel complex typically occupies most of the valley floor, leaving little room for a floodplain. Glacial outwash streams and ephemeral streams draining mountainous areas in arid regions are normally braided, and may form broad sheets of sand or gravel crossed by networks of shallow, shifting channels.

Meandering rivers. These are single-channel streams of high sinuosity, in which islands and midchannel bars are rare. Sediment in these rivers range from very coarse to very fine. A significant proportion of the bedload typically is deposited on the insides of meander bends, forming point bars. The channel, with its coarse deposits, may be confined to a narrow belt within an alluvial valley, flanked by a broad floodplain, upon which deposition of fine-grained sediment takes place only during flood events—seasonally or at longer intervals.

Anastomosed rivers. These develop in stable, low-energy environments or in areas undergoing rapid aggradation. They consist of a network of relatively stable, low- to high-sinuosity channels bounded by well-developed floodplains. Channels are characteristically narrow and accumulate narrow, ribbonlike sandstone bodies.

Straight channels. These are rare, occurring mainly as distributaries in some deltas.

Rivers which emerge from a mountainous catchment area into a low plain drop their sediment load rapidly. The channel may bifurcate, becoming braided in character. A distinctive landform, an alluvial fan, results.

Sedimentary facies. River deposits of sediment occur as four main types.

Channel-floor sediments. The coarsest bedload is transported at the base of the channel, commonly resulting in deposits of gravel (**Fig. 1***a*), waterlogged vegetation, or fragments of caved bank material. In

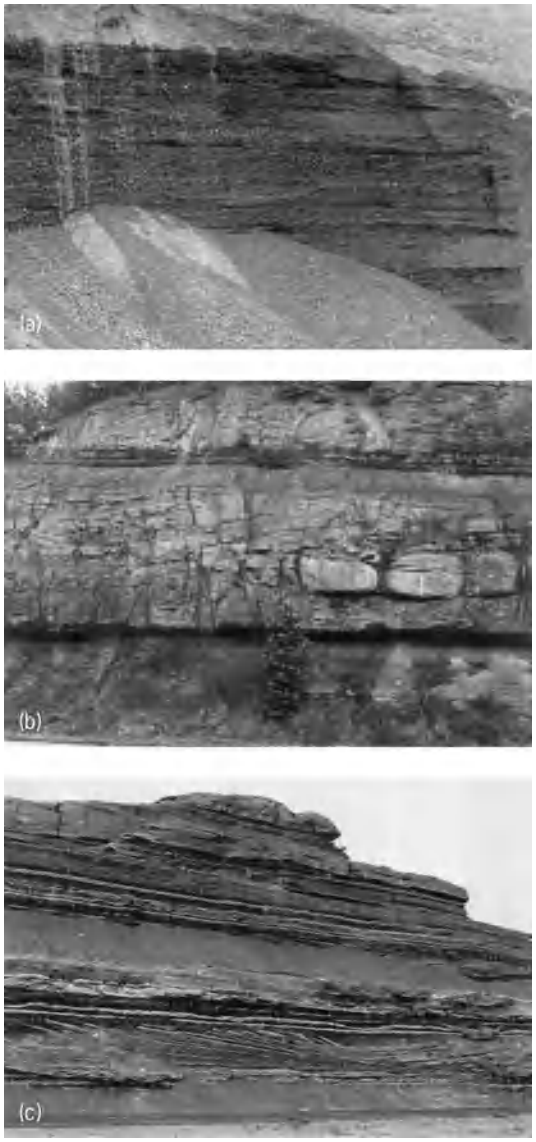


Fig. 1. Typical fluvial deposits. (*a*) Gravel and sand channel-fill and bar deposits exposed in a gravel quarry face about 40 ft (12 m) high, fluvioglacial outwash, Alberta, Canada. (*b*) Point bar, 13 ft (4 m) thick, Carboniferous, Alabama. (*c*) Typical floodplain deposits, Triassic, Arizona; outcrop is about 33 ft (10 m) high.

sand-bed rivers the channel floor commonly is covered by fields of large, sinuous-crested dunes or ripples (with amplitudes of 2 in. to 10 ft, or 5 cm to 3 m), which impart a trough–cross-bedded structure to the sand.

Bar sediments. Accumulations of gravel, sand, or silt occur along river banks and are deposited within channels, forming bars that may be of temporary duration, or may last for many years, eventually becoming vegetated and semipermanent. Bars attached to one of the channel banks are termed side or lateral bars. Those occurring on the insides of meander bends are termed point bars. They develop by lateral accretion as the meander widens or shifts in position by erosion on the outer bank of the bend (Fig. 1*b* and **Fig. 2**). Bars occurring within channels accrete by the addition of sediment on all sides, but most commonly preferentially on one side representing the inside of a bend in the adjacent channel, or at the downstream end of the bar. Such bars commonly have complex internal structures, reflecting many seasons of growth and intervals of erosion.

Channel-top and bar-top sediments. These are typically composed of fine-grained sand and silt, and are formed in the shallow-water regions on top of bars, in the shallows at the edges of channels, and in abandoned channels. Small-scale ripples, with amplitudes of less than 2 in. (5 cm), are typical sedimentary structures, together with roots and bioturbation structures.

Floodplain deposits. These are formed when the water level rises above the confines of the channel and overflows the banks (Fig. 1*c*). Much of the coarser sediment is deposited close to the channel, in the form of levees. Breaks in the channel bank, termed crevasses, permit the transportation of additional coarse sediment onto the floodplain, where it forms small deltalike bodies spreading out into the floodplain, termed crevasse splays. Much silt and mud may be carried considerable distances from the channel, forming blanketlike deposits. In swampy areas, floodplains may be the site of thick vegetation, which in time may be transformed into lignite and eventually into coal. Soils develop in response to weathering activity and plant growth, and may form distinctive brightly colored layers, termed paleosols. Nodular beds of calcium carbonate are a common component of paleosols, especially in relatively arid areas, where they form as a result of the evaporation of ground waters. *See* PALEOSOL.

Facies associations and sedimentary cycles. Fluvial sediments may be dominantly conglomeratic, sandy, or silty, depending on the nature of the sediment load of the river. This characteristic is mainly a function of slope and the proximity to sources, but is also a reflection of sediment availability. Certain source materials, such as fine-grained sediments or limestones, may yield coarse debris on erosion, but it

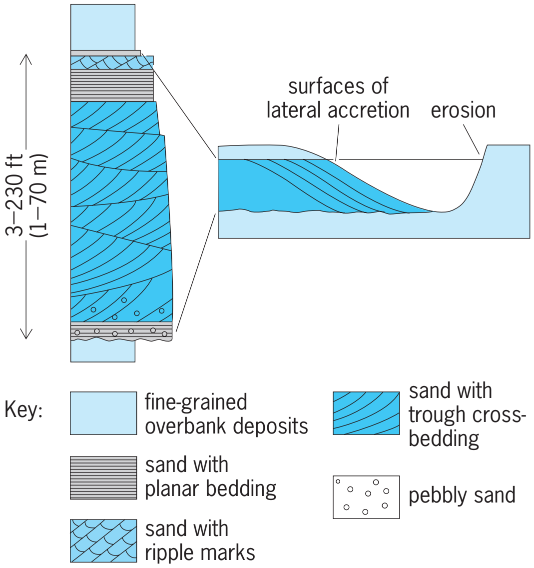


Fig. 2. Development of a fining-upward succession by lateral accretion of a point bar, such as that in Fig. 1*b*.

is likely to be broken down into fine material or dissolved on prolonged transportation. Humid climates favor chemical and biochemical weathering processes, which yield a large suspended or dissolved sediment load. Coarse detritus is more typically the product of drier climates, in which mechanical weathering processes (such as frost shattering) are dominant.

The sedimentary facies described above were listed in approximate vertical spatial order, from channel floor to floodplain. This order is one of decreasing grain size upward, a feature which may commonly be observed in ancient fluvial deposits (Fig. 2). Such deposits may consist of a series of fining-upward successions, or cycles, each a few meters to few tens of meters in thickness.

There are a variety of causes of such cycles. The first that was recognized is the mechanism of lateral accretion, whereby point bars enlarge themselves in a horizontal direction as the meander bounding them migrates by undercutting the bank on the outside of the bend (Fig. 2). The depositional surface of the point bar may be preserved as a form of large-scale, low-angle cross-bedding within the deposit, its amplitude corresponding approximately to the depth of the channel. Similar cycles may be caused by the nucleation and growth of large compound bars or sand flats within braided channel systems. The accretion surfaces, in such cases, may dip in across- or down-channel directions. Individual flood events, especially on the sand flats of ephemeral stream systems, may form sheetlike flood cycle deposits up to a meter or so thick, the upward fining corresponding to decreasing energy levels as the flood waned. The gradual choking of a channel with sediment, and the progressive abandonment of the channel, will also generate a fining-upward cycle.

Regional controls. Tectonic activity in a fluvial catchment area may cause the generation of sedimentary cycles, which either fine or coarsen upward, depending on whether relief and slope are decreased or increased, respectively. Such cycles tend to be tens to hundreds of meters thick and to extend for several to many kilometers. They may have smaller cycles formed by channel fill and migration processes nested within them.

River systems are also affected by changes in base level, that is, by a rise or fall in the level of the lake or sea into which the river drains. A fall in base level may lead to widespread incision of channels along a coastal plain as they adjust to a lower river mouth. Between channels, sedimentation may cease, with the formation of widespread, well-developed paleosols. The same effect is brought about by peneplanation, that is, long-continued subaerial erosion in the absence of tectonic rejuvenation of the river system.

A rise in base level may flood the mouths of rivers, forming estuaries. However, if the sediment supply is adequate, sedimentation may be able to keep pace with such a base-level rise, with the river changing in style in response to changes in the balance between the rate of sediment input and the rate of base-level change (**Fig. 3**). These sedimentary responses to external forcing are part of a larger story concerning the regional and global controls of sedimentation. The regional erosion surface formed at a time of falling to low base level or deep peneplanation is termed a sequence boundary. It may be cut by incised channels that are, in turn, typically filled by coarse channel sediments (in sequence stratigraphic terminology, these deposits are classified as the lowstand systems tract). At the coast a rise in base level is commonly recorded by transgression (the transgressive systems tract). The coastal rivers at this time may be of anastomosed style. As base-level rise reaches its highest level, the rate of rise slows, and the rivers typically evolve into a meandering type (highstand systems tract). These changes may be reflected in the resulting sediments by changes in the geometry and spacing of channel sandstone and conglomerate bodies. Lowstand deposits commonly consist of coarse, laterally amalgamated channel deposits. A rapid rise in base level (the transgressive systems tract) may be marked by isolated channel sands spaced out within thick floodplain units. The spacing of such channel bodies becomes closer, and more units are amalgamated, in the highstand deposits, above. *See* FLOODPLAIN; SEQUENCE STRATIGRAPHY.

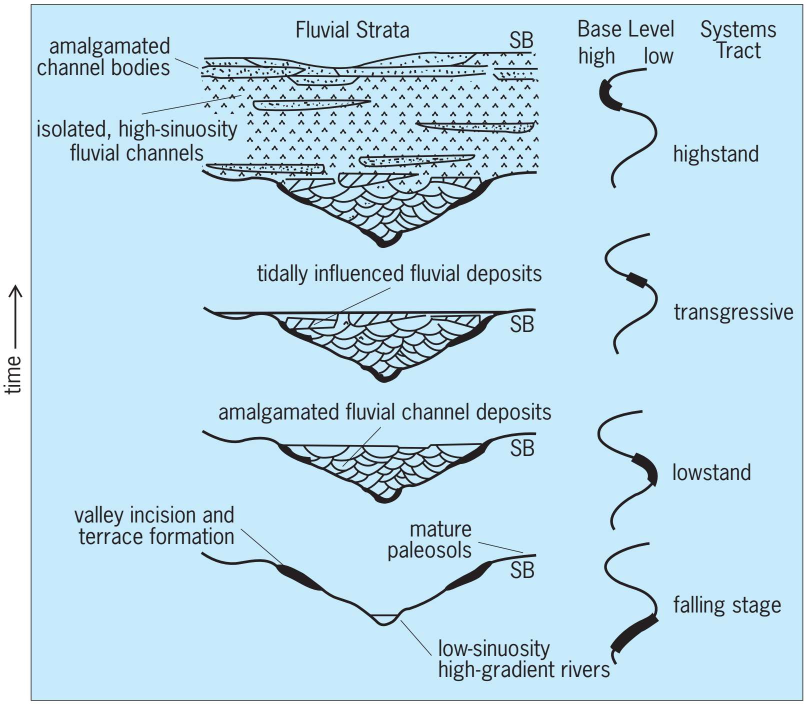


Fig. 3. Response of river systems and their deposits to a cycle of fall and rise of base level. SB = sequence boundary. (*Modified from K. W. Shanley and P. J. McCabe, Perspectives on the Sequence Stratigraphy of Continental Strata, American Association of Petroleum Geologists, vol. 78, p. 560, 1994*)

Far inland, base-level changes may not markedly affect the rivers unless they persist for very long periods of time. Changes in discharge and sediment yield in response to climate change are commonly more important in controlling river style and the resulting sediment types. For example, during the Pleistocene glaciation, rapid deposition of coarse sediments occurred at the margins of continental ice caps.

Channel erosion, forming widespread surfaces of incision, occurred at times of change, between glacial and interglacial periods, because at these times river discharge tended to increase whereas sediment yield did not. Climatically driven erosion and deposition inland were therefore out of phase with the cycle of change generated by base-level change at the coast.

More than one forcing function, including climate change, base-level change, and tectonism, may be operating at any one time, resulting in complex patterns of cyclicity. The resulting sequences may be widespread. Reconstruction of this sequence stratigraphy may provide an essential mapping tool for those engaged in basinal exploration.

Tectonic setting of fluvial deposits. The thickest (up to 6 mi or 10 km) and most extensive fluvial deposits occur in convergent plate-tectonic settings, including regions of plate collision, because this is where the highest surface relief and consequently the most energetic rivers and most abundant debris are present. Some of the most important accumulations occur in foreland basins, which are formed where the continental margin is depressed by the mass of thickened crust formed by convergent tectonism. Examples include the modern Himalayan foreland basin of the Indus and Ganges valleys, the Devonian foreland basin west of the Appalachian Mountains, and the late Cenozoic foreland basin of France and Germany, north of the alpine mountain chain. *See* BASIN.

Thick fluvial deposits also occur in rift basins, where continents are undergoing stretching and separation. The famous hominid-bearing sediments of Olduvai Gorge and Lake Rudolf are fluvial and lacustrine deposits formed in the East Africa Rift System. Triassic fault-bounded basins along the North American Atlantic coast and through western Europe are an older but comparable example. Fluvial deposits are also common in wrench-fault basins, such as those in California.

Economic importance. Significant volumes of oil and gas are trapped in fluvial sandstones. Major reservoirs include those of Triassic-Jurassic age in the North Sea Basin; Triassic sandstones of the Paris Basin; Permian-Triassic sandstones of Prudhoe Bay on the Alaskan North Slope; the Lower Cretaceous reservoirs of the giant Daqing field of the Songliao Basin, China; Jurassic sandstones of interior Australia; the heavy-oil sands of the Cretaceous Athabasca and related deposits in Alberta; and numerous large to small fields in mature areas such as the Alberta Basin (Cretaceous), the southern midcontinent (Pennsylvanian), and the Gulf Coast (Cretaceous and Oligocene fields).

Placer gold, uranium, and diamond deposits of considerable economic importance occur in the ancient rock record in South Africa and Ontario, Canada, and in Quaternary deposits in California and Yukon Territory. Economically significant roll-front uranium deposits occur in the Mesozoic deposits of the American Western Interior and elsewhere, primarily in fluvial facies.

Fluvial deposits are also essential aquifers, especially the postglacial valley-fill complexes of urban Europe and North America. Much work needs to be done to investigate the internal geometry of these deposits in order to resolve problems of domestic and industrial pollution that now interfere with the use of the ground water from these sources.

Andrew D. Miall

Bibliography. P. A. Carling and M. R. Dawson (eds.), *Advances in Fluvial Dynamics and Stratigraphy*, Wiley, 1996; C. R. Fielding (ed.), Current research in fluvial sedimentology, *Sed. Geol.*, special issue, vol. 85, 1993; A. D. Miall, *The Geology of Fluvial Deposits*, Springer-Verlag, 1996.

Fluxional compounds

Molecules that undergo rapid intramolecular rearrangements among equivalent structures in which the component atoms are interchanged. The rearrangement process is usually detected by nuclear magnetic resonance (NMR) spectroscopy, which can measure rearrangement rates from 0.5 to 10,000 s^{-1}. With sufficiently rapid rates, a single resonance is observed in the NMR spectrum for a molecule that might be expected to have several nonequivalent nuclei on the basis of its instantaneous structure.

Organic structures. Within organic chemistry, degenerate Cope rearrangements represented some of the first examples of interconversions between equivalent structures, but these were relatively slow (**Fig. 1***a*). The rate of this rearrangement is rapid in more complex molecules. The epitome of degeneracy is reached in bullvalene, which has more than 1,200,000 equivalent structures (Fig. 1*b*) and rapidly interconverts among them.

Fluxional molecules are frequently encountered in organometallic chemistry, and rapid rearrangements which involve migrations about unsaturated organic rings are commonly observed. The best known (called ring-whizzers) are cyclopentadienyl and cyclooctatetraene complexes of iron (**Fig. 2**).

Inorganic structures. Inorganic structures also exhibit fluxional phenomena, and five-coordinate complexes provide the greatest number of well-known

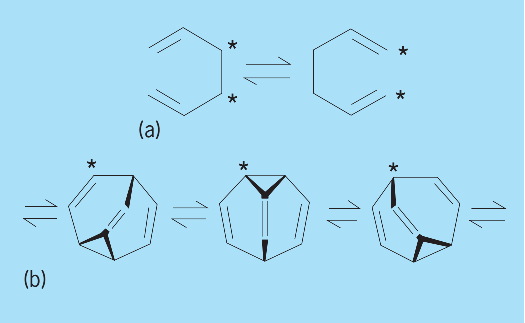


Fig. 1. Degenerate rearrangements. (*a*) Cope rearrangement. (*b*) Rearrangement of bullvalene. The asterisks indicate labels for the atoms so that they can be distinguished from one another.

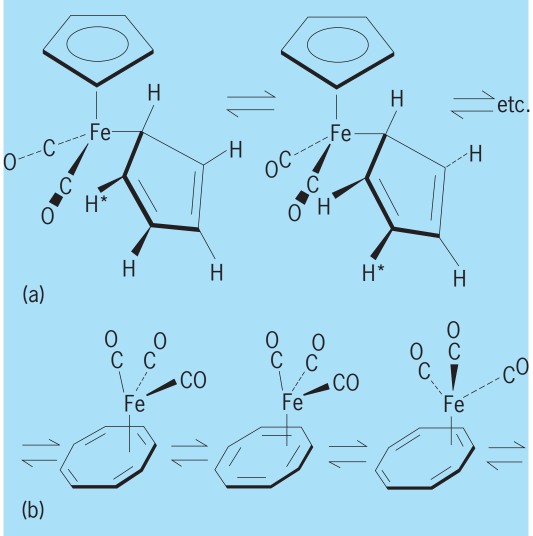


Fig. 2. Rearrangements about unsaturated organic rings. (*a*) Cyclopentadienyl iron complex. (*b*) Cyclooctate-traene iron complex.

examples. Although the static structure would suggest a significant difference between equatorial and axial fluorine nuclei, phosphorus pentafluoride rearranges rapidly and shows only a single ^{19}F resonance split by P-F spin-spin coupling even though the static structure would suggest a significant difference between equatorial and axial fluorines. The retention of this coupling during the rearrangement demands that the process be intramolecular. This distinguishes a fluxional process from other intermolecular exchange processes which might give rise to dynamic NMR spectra. Detailed line-shape analysis of the NMR spectra can also provide insight into the possible paths of the rearrangements. Thus, the ring-whizzers have been shown to rearrange preferentially by moving to an adjacent position rather than across the ring.

The rearrangement of PF_5 involves interconversion of the trigonal bipyramidal molecule (**1**) to a square pyramidal configuration (**2**) and back (**Fig. 3**). The square pyramidal structure is of sufficiently high energy that none of that structure is observed by spectroscopic techniques that operate on a faster time scale than does NMR. If two such nonequivalent structures are present in observable concentrations but interconvert rapidly to cause averaging in the NMR experiments, they are said to be stereochemically nonrigid. This term is generally taken to embrace all compounds that undergo rapid reversible intramolecular rearrangements. Thus, fluxional compounds are a subset of nonrigid compounds with equivalent structures. Nonequivalent structures, that is, tautomers, might be stereochemically nonrigid if they rearranged rapidly, but would not be considered fluxional.

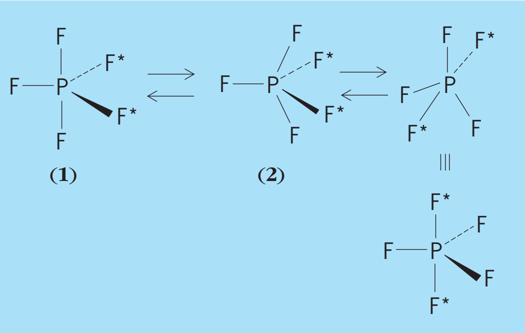


Fig. 3. Rearrangement of PF_5. Structure (2) interconverts to the equivalent structure on the right.

Some workers prefer to reserve the term fluxional for molecules in which bonds are broken and reformed in the rearrangement process. Hence, of the examples above, only bullvalene and the iron complexes would be termed fluxional, whereas all would be considered stereochemically nonrigid. *See* NUCLEAR MAGNETIC RESONANCE (NMR); RESONANCE (MOLECULAR STRUCTURE); TAUTOMERISM.

J. W. Faller

Bibliography. C. J. Ballhausen, *Molecular Electronic Structures of Transition Metal Complexes*, 1979; F. A. Cotton, Fluxional organometallic molecules, *Acc. Chem. Res.*, 1:257–265, 1968; W. von E. Doering and W. R. Roth, Thermal rearrangements, *Angew. Chem.* (Internat. Ed.), 2(3):115–122, 1963; J. W. Faller, Fluxional and nonrigid behavior of transition metal organometallic π-complexes, *Adv. Organometal. Chem.*, 16:211–239, 1977; L. M. Jackman and F. A. Cotton, *Dynamic Nuclear Magnetic Resonance Spectroscopy*, 1975.

Flywheel

A rotating mass used to maintain the speed of a machine between given limits while the machine releases or receives energy at a varying rate. A flywheel is an energy storage device. It stores energy as its speed increases and gives up energy as the speed decreases. The specifications of the machine usually determine the allowable range of speed and the required energy interchange.

Theory. The energy, speed, and size of a flywheel are related by Eq. (1) for use with SI units, and by Eq. (2) for American-British engineering units. The term $n^2{}_1 - n^2{}_2$ may be replaced by the product $2c_f n^2$,

$$Mk^2 = \frac{2E}{n_1 - n_2^2} \tag{1}$$

$$Wk^2 = \frac{182.4gE}{n_1^2 - n_2^2} \tag{2}$$

where W = flywheel weight (lb)
M = flywheel mass (kg)
E = energy change (joules or ft · lb)
k = radius of gyration of flywheel (m or ft)
n_1, n_2 = maximum and minimum velocity (angular) of flywheel (radians/s or revolutions/min)
g = gravity constant (32.174 ft/s^2)

which gives Eqs. (3) and (4). The coefficient of fluctuation is given by Eq. (5), and the average speed or angular velocity is given by Eq. (6).

$$Mk^2 = \frac{E}{c_f n^2} \tag{3}$$

$$Wk^2 = \frac{91.2gE}{c_f n^2} \tag{4}$$

$$c_f = \frac{\text{coefficient of}}{\text{fluctuation}} = \frac{n_1 - n_2}{n} \tag{5}$$

$$n = \frac{\text{average speed}}{\text{or angular velocity}} = \frac{n_1 + n_2}{2} \tag{6}$$

The energy interchange cycle of a flywheel may occupy any amount of angular motion about its center of rotation. Most frequently it is a cycle of one or more revolutions, but in internal combustion engines with multiple cyclinders, it may be only a small part of a revolution.

All rotating components of a machine contribute to the flywheel effect and should be considered in a complete analysis of the design. The term Mk^2 or Wk^2 in Eqs. (1)–(4) is the total flywheel effect of all rotating components and is made up of the sum of the Wk^2 or Mk^2 for all such components as shown in Eqs. (7) and (8), where subscript f refers to the flywheel and subscripts 1 and 2, to the other rotating components.

$$M_1k_1^2 + M_2k_2^2 + \cdots + M_fk_f^2 = \sum M_1k_1^2 + M_fk_f^2 \tag{7}$$

$$W_1k_1^2 + W_2k_2^2 + \cdots + M_fk_f^2 = \sum Wk^2 + M_fk_f^2 \tag{8}$$

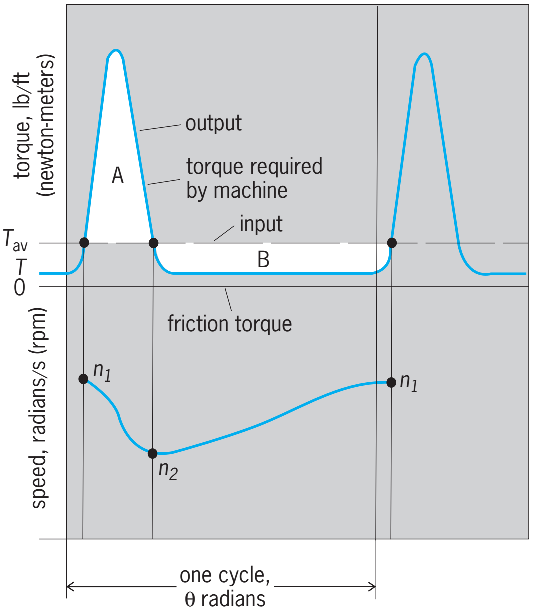


Fig. 1. Determination of energy change for a flywheel for a typical power-press torque requirement.

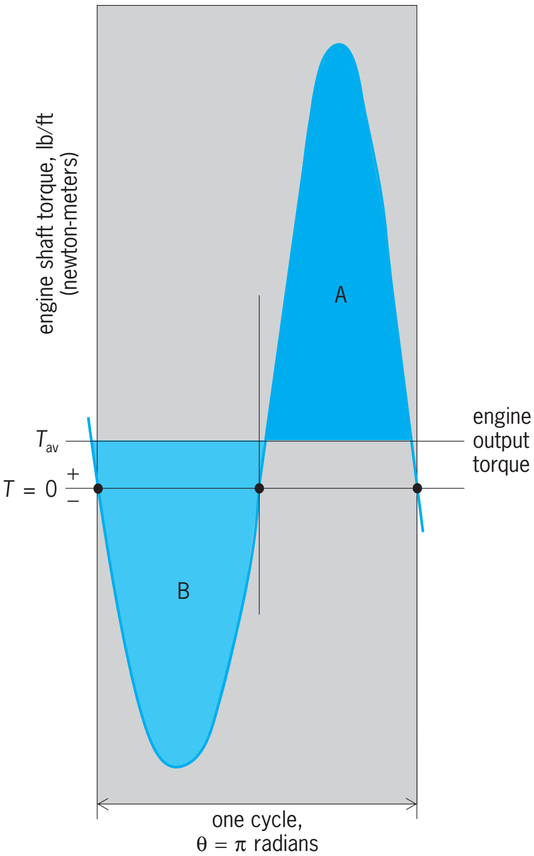


Fig. 2. Determination of energy change for a flywheel for a typical four-cylinder, four-cycle internal-combustion-engine torque requirement.

Equations (3) and (4) can now be written in more useful terms for the designer as Eqs. (9) and (10).

$$M_fk_f^2 = \frac{E}{c_f n^2} - \sum Mk^2 \tag{9}$$

$$W_fk_f^2 = \frac{91.2gE}{c_f n^2} - \sum Wk^2 \tag{10}$$

Flywheel design. In the design of a flywheel the first step is to determine the energy interchanged. This can be done best by plotting the torque required by the machine against the angular motion of the flywheel for one cycle of θ radians (**Figs. 1** and **2**). In this plot, the area under the torque curve represents the total energy requirements of the machine during the cycle. Then, average torque T_{av} defined as the total energy of the cycle E_{tot} divided by the length of the cycle θ in radians is drawn on the same torque scale. The areas between the original torque curve and the average torque curve above and below this average torque (shaded areas in Figs. 1 and 2) should be equal and are the E, or change in energy, in Eqs. (1) and (2) or (9) and (10).

The second step is to establish the speed range n_1 and n_2 of the machine or to select the appropriate coefficient of fluctuation from the **table**.

The third step is to calculate the flywheel effect $M_fk^2{}_f$ or $W_fk^2{}_f$ with Eqs. (9) and (10), calculating

Coefficients of fluctuation

Machine	C_f
Crushers, hammers, punches, shears	0.100–0.200
Machine tools	0.020–0.030
Electric generators	0.003–0.005
Compressors, pumps, and engines	0.040–0.050
Textile machinery	0.010–0.030

ΣMk^2 or ΣWk^2 of all the significant rotating components. The value of c_f is taken from the table or calculated by Eqs. (5) and (6). The value of E is established from Figs. 1 and 2: area A = area B = energy change E. The angular velocity (rotational speed) should come from the machine specifications. Having thus established numerical values for $M_f k^2_f$ or $W_f k^2_f$, the designer is now in a position to calculate the diameter and width of the flywheel to meet the flywheel effect requirements. The designer must vary all the variables that affect mass or weight to fit the space, speed, and safety requirements of the machine.

Flywheel construction. The difficulty of casting stress-free spoked flywheels leads the modern designer to use solid web castings or welded structural steel assemblies. For large, slow-turning flywheels on heavy-duty diesel engines or large mechanical presses, cast spoked flywheels of two-piece design are standard (**Fig. 3**).

Because it is difficult to calculate the stresses in a spoked flywheel, these are usually designed on the basis of a maximum allowable peripheral speed of 5000–7000 ft/min (1500–2100 m/min) for cast iron. The higher value is for rims cast by special techniques to eliminate rim porosity or blowholes. Higher rim speeds to 10,000 ft/min (3000 m/min) may be used with cast steel. Structural steel plate welded flywheels of the solid web type can be accurately analyzed for stress by applying the equations of rotating disks to the components of the flywheel and allowing a factor for the stresses caused by welding.

Unusual flywheel applications. Flywheels of appropriate design can be used to store relatively large quantities of energy for use at a later time away from their spin-driving energy source. To store large quantities of energy in a reasonably small space requires flywheel designs that can spin at very high speeds up to 20,000–30,000 revolutions per minute. They must be enclosed in a near-perfect vacuum to reduce air friction losses and be equipped with very high-efficiency transmission systems. These high-speed flywheels must also be designed with extremely high-tensile-strength materials to withstand the stresses produced at their maximum speed. *See* ANGULAR MOMENTUM; ENERGY; INERTIA; MASS; MOMENT OF INERTIA; MOMENTUM.

L. Sigfred Linderoth

Bibliography. R. C. Johnson, J. E. Shigley and C. R. Mischke, *Mechanisms: A Mechanical Designer's Workbook*, 1990; A. Parrish (ed.), *Mechanical Engineer's Reference Book*, 12th ed., 1994.

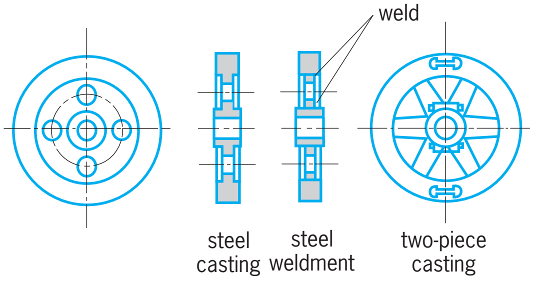


Fig. 3. Typical flywheel structures.

Foam

A material composed of gas bubbles separated from one another by liquid or solid sheets (or lamellae). Foam densities range from nearly zero to about 43 lb/ft^3 (0.7 g/cm^3), above which gas emulsions rather than foams are formed. The bubbles are spherical when the lamellae that separate the bubbles are thick (approximately 0.04 in. or 1 mm); the bubbles are polyhedral when the lamellae are thin (approximately 0.0004 in. or 0.01 mm). Pure liquids do not foam; that is, they cannot produce liquid lamellae of any permanence. Relatively permanent lamellae are created only when a substance is present that is adsorbed at the surface of the liquid. Such substances may be in true solution in the liquid, or they may be particles of a finely divided solid which, because of poor wetting by the liquid, remain at the surface. In both cases, surface layers rich in the added substance are formed. The inability of the adsorbed substance to enter the bulk of the liquid preserves the surface and so confers stability on the foam lamellae. However, this stability is transient. Gravity drains solution out of the lamellae, and the solubility of gas in the solution, although slight, allows gas to diffuse out of the foam. These destabilizing factors tend to separate the liquid and the gas phases of the foam and so destroy it. *See* ADSORPTION; EMULSION; INTERFACE OF PHASES.

On the other hand, the stability of the foam is enhanced if the adsorbed layer is thick and gelatinous. These attributes confer two stabilizing properties on the liquid lamellae: resilience and plasticity. *See* SURFACTANT.

Resilience and plasticity of lamellae. The resilience, or elastic recovery after externally induced mechanical stress, of a liquid lamella originates in the higher surface tension of a freshly formed surface of any solution that contains a surface-active solute. This occurs because such a surface contains relatively few solute molecules. Then, as solute molecules are adsorbed, the surface tension declines. The higher dynamic surface tension of a nascent surface, if it lasts long enough, can heal any incipient breach by drawing toward it the surrounding surface of lower surface tension. Drawn along with it, by viscous drag, is a considerable amount of subsurface solution. This restoring mechanism depends on the liquid lamella being sufficiently thick to provide enough subsurface solution. Prolonged drainage impairs resilience because of extreme depletion of subsurface solution. Aged lamellae, defective in this healing mechanism, are brittle. A lack of resilience is the cause of the ultimate collapse of foams. *See* SURFACE TENSION.

An additional factor, plasticity of the surface, can prolong stability from mere seconds to several hours. Solutions with this latter property are perfectly fluid in their flow properties as bulk liquids, but the surface acts as if it were a thin sheet of a soft plastic material. The plastic layer has a depth of 10–40 micrometers, so that a lamella, which is plastic at both surfaces, may contain very little interstitial solution that is free to drain out. The lamellae become totally plastic, and the foam is correspondingly long-lived. The presence of water-soluble proteins, such as casein or albumin, provides stability in foams produced with heavy cream, egg white, beer, and natural rubber latex. Surface plasticity adds enormously to the stability of foams, as exemplified in meringue, whipped cream, fire-fighting foams, and shaving foams. Finely divided solids may have the same effect, as exemplified in the industrial process of mineral separation by means of froth flotation.

Resilience and plasticity of lamellae are two distinct properties, corresponding to the observation that there are two distinct types of foams, which differ considerably in stability. Ephemeral or transient foams (such as champagne foam) arise from resilience alone. Such foams drain rapidly to a critical thickness of 20–30 nanometers, at which point the lamellae are brittle and soon rupture. The very stable foams (such as beer foam) begin in the same way, and instantaneous rupture is prevented by resilience. But this effect is soon replaced by the slow growth of gelatinous surface layers, which serve to preserve the lamellae. At the extreme, the lamellae may be stabilized permanently by polymerization of the medium.

Plastic foams. Plastic foams (also known as cellular, expanded, or sponge plastics) are solid foams in which the gas can be considered as a lightweight filler or extender of the plastic phase. Plastics commonly used are polyurethane, synthetic rubber, polystyrene, polyvinyl chloride, and polyolefins. In all cases the foam is produced while the medium, a monomer or a blend of monomers, is still fluid. Polymerization of the medium occurs before the foam has had time to collapse substantially. The nature of the polymer and the foam density are dominant factors in determining the physical properties of cellular plastics. Practical foam densities may vary from about 0.5 to 60 lb/ft^3 (0.008 to 0.96 g/cm^3) or almost up to the density of the solid plastic. If the foam cells are interconnected so that gas can pass freely among them, the foam is termed open-cell; if the gas is sealed in each cell, the foam is termed closed-cell. Flexible foams are usually open-cell; rigid foams are usually closed-cell. The major applications of foamed plastics are for insulation and packaging (rigid foams) and comfort products (flexible foams). *See* POLYMER.

Elimination. Although foams of exceptional stability are desired in some industrial applications, foam is a nuisance in many other situations. In the glue and paper industries, in the processing of foodstuffs, in sewage disposal, in the dehydration of oil, and in distillation and fractionation towers, foams cause undesirable complications. A common recourse is to add chemical antifoam agents, which are usually insoluble liquids of very low surface tension. When a droplet of such a liquid is sprayed onto the foam or is carried into it by mechanical agitation, it behaves as a local area that resists wetting by the medium. This results in a mechanical action, that is, the retraction of the lamella from the insoluble droplet situated within it; the process is known as dewetting. The force of its mechanical action— or, as some conceive it, the bridging of the liquid lamella by a droplet whose adhesion to the medium is too weak to hold the composite structure together—causes it to collapse. *See* COLLOID. Sydney Ross

Bibliography. K. C. Frisch and J. H. Saunders (eds.), *Plastic Foams*, 1972–1976; D. Myers, *Surfactant Science and Technology*, 2d ed., 1992; S. Ross and I. D. Morrison, *Colloidal Systems and Interfaces*, 1988.

Focal length

A measure of the collecting or diverging power of a lens or an optical system. Focal length, usually designated f' in formulas, is measured by the distance of the focal point (the point where the image of a parallel entering bundle of light rays is formed) from the lens, or more exactly by the distance from the principal point to the focal point. *See* GEOMETRICAL OPTICS.

The power of a lens system is equal to n'/f', where n' is the refractive index in the image space (n' is usually equal to unity). A lens of zero power is said to be afocal. Telescopes are afocal lens systems. *See* DIOPTER; LENS (OPTICS); TELESCOPE. Max Herzberger

Fog

A cloud comprising waterdroplets or (less commonly) ice crystals formed near the ground and resulting in a reduction in visibility to below 0.6 mi (1 km). This is lower than that occurring in mist, comprising lower concentration of waterdroplets, and haze, comprising smaller-diameter aerosol particles.

Composition. Fog results from the cooling of moist air below its saturation (dew) point. Droplets form on hygroscopic nuclei originating from ocean spray, combustion, or reactions involving trace chemicals in the atmosphere. Visibility is reduced even more when such nuclei are present in high concentrations and faster cooling rates activate a larger fraction of such nuclei. Thus, polluted fog, with more numerous smaller droplets, results in lower visibility for a given water content. *See* DEW POINT.

Haze, the precursor to fog and mist, forms at relative humidity below 100% to about 80%. It is composed of hygroscopic aerosol particles grown by absorption of water vapor to a diameter of about 0.5 micrometer, concentration 1000 to 10,000 per cubic centimeter. Fog and mist form as the relative humidity increases just beyond saturation (100%), so that larger haze particles grow into cloud droplets with a diameter of 10 μm and a concentration of several hundred per cubic centimeter. Fog and mist are a mix of lower-concentration cloud droplets and higher-concentration haze particles. By contrast, smog is formed of particles of 0.5–1-μm diameter, produced by photochemical reactions with organic vapors from automobile exhaust. *See* ATMOSPHERIC CHEMISTRY; HUMIDITY; SMOG.

Formation. The saturated vapor pressure of water declines with falling temperature; as air containing water vapor is cooled below its dew point, excess vapor condenses to droplets. Fog forms in this way by cooling that results from several processes.

Radiation fog forms under near calm and clear nighttime conditions by cooling, which results from radiation loss to the cold sky. Such fog often appears first as shallow layers a few feet thick just above moist ground (**illus.** *a*). Both the surface of the ground and moist air cool by radiation, resulting in vapor deposits as fog droplets above the ground and, very close to the surface, vapor deposits on the ground or on vegetation as dew. Subsequently, slight air motion, resulting from the settling of denser, colder air results in local lifting of some air, with fog formation at a slightly higher level following expansion cooling. This enables the fog to build upward and reduces radiation cooling at lower levels. Such motions can also lead to fog formation by the mixing of saturated air parcels at different temperatures (analogous to a person's breath on a cold morning), leading to the phenomenon of "Arctic frost" or sea smoke, that is, shallow wispy layers of fog just above a warmer water surface. *See* DEW.

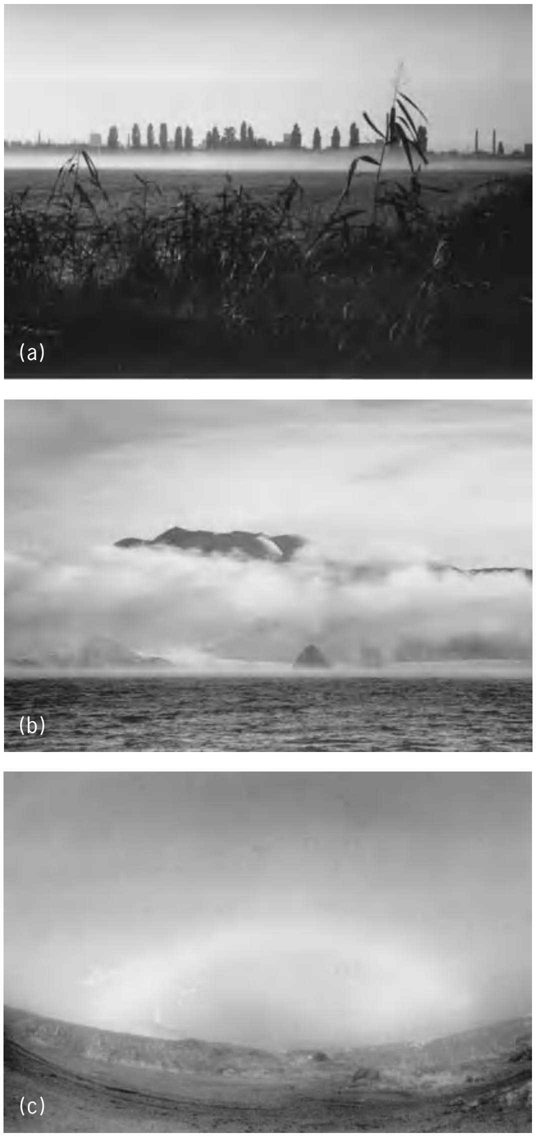
Advection fog forms when moist air moves over a colder surface, such as snow or hail that has recently fallen, or cold water, as near Newfoundland. Both radiation and advection fog build to form deeper layers of lifted fog, which are then transformed to stratocumulus clouds (illus. *b*) when mixed by increasing winds aloft or by solar heating as the Sun rises. This process ultimately leads to "burn off" of the fog over land and low clouds. *See* CLOUD.

Hill fog forms when air flows over high ground and cools by expansion; it is in many ways similar to clouds formed by convective processes.

Supercooled fog forms at temperatures below 32°F (0°C). This fog may be cleared over a localized area by artificial cloud seeding, providing the temperature is below about 25°F (−4°C). *See* WEATHER MODIFICATION.

At temperatures below about −31 to −40°F (−35 to −40°C), all waterdroplets freeze and an ice fog forms. This is particularly troublesome in communities experiencing such temperatures, because as the temperature falls, nuclei and moisture produced by combustion form water fog first and then ice fog.

Deep fogs may give a light drizzle. Viewed from above, the presence of droplets of intermediate size (50 μm) that coalesce to drizzle is seen as a fog bow, a white, almost colorless arc at the rainbow angle (42°) [illus. *c*]. *See* CLOUD PHYSICS; RAINBOW.



Fogs formed by various processes. (*a*) Radiation fog; appears first as shallow layers above moist ground. (*b*) Stratocumulus cloud, formed from lifted fog. (*c*) Fog bow resulting from large cloud droplets in the fog top.

John Hallett

Bibliography. F. K. Lutgens and E. J. Tarbuck, *The Atmosphere: An Introduction to Meteorology*, 6th ed., 1995; S. Twomey, *Atmospheric Aerosols*, 1977.

Fold and fold systems

Folds are recognized where layered rocks have been distorted into wavelike forms. Some folds are fractions of an inch across and have lengths measured in inches, whereas others are a few miles wide and tens of miles long.

Elements of folds. The axial surface (**Fig. 1**) divides a fold into two symmetrical parts, and the intersection of the axial surface with any bed is an axis. In

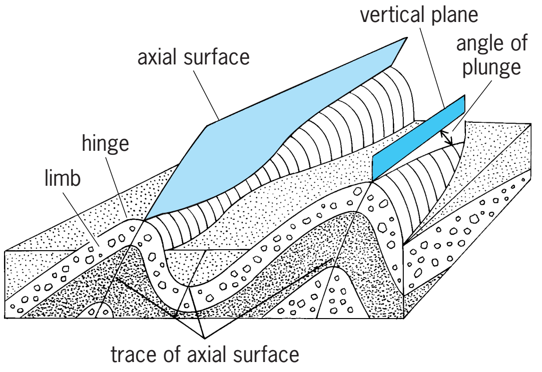


Fig. 1. Elements of folds.

general, an axis is undulatory, its height changing along the trend of the fold. Relatively high points on an axis are culminations; low points are depressions. The plunge of a fold is the angle between an axis and its horizontal projection. The limbs or flanks are the sides. A limb extends from the axial surface of one fold to the axial surface of the adjacent fold. Generally, the radius of curvature of a fold is small compared to its wavelength and amplitude, so that much of the limb is planar. The region of curvature is the hinge. *See* ANTICLINE; SYNCLINE.

Geometry of folds. The geometry of folds is described by the inclination of their axial surfaces and their plunges (**Fig. 2**). Upright folds have axial surfaces that dip from 81° to 90°; inclined folds have axial surfaces that dip from 10° to 80°; and recumbent folds have axial surfaces that dip less than 10°. Vertical folds plunge from 81° to 90°; plunging folds plunge from 10° to 80°; and horizontal folds plunge less than 10°. Auxiliary descriptive terms depend on the attitude or the relative lengths of the limbs. Overturned folds are inclined folds in which both limbs dip in the same direction; isoclinal folds are those in which both limbs are parallel; symmetrical folds have limbs of equal length; and asymmetrical folds have limbs of unequal length. The descriptions of folds consist of combinations of the above terms, for example, isoclinal upright horizontal fold, overturned plunging fold, asymmetrical inclined horizontal fold.

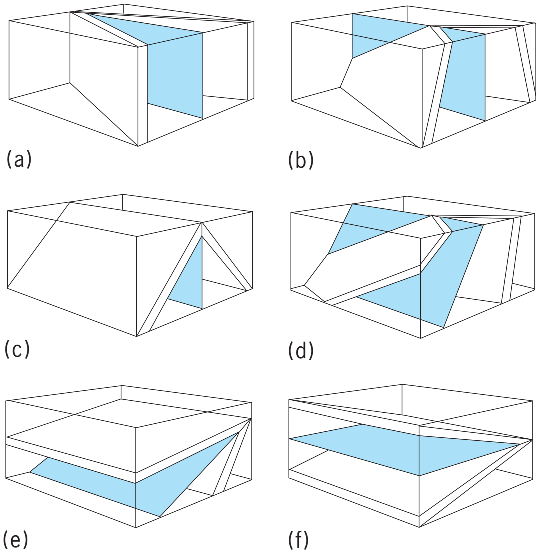


Fig. 2. Description of folds based on the attitudes of their axial surfaces and their plunges. (*a*) Vertical. (*b*) Upright plunging. (*c*) Upright horizontal. (*d*) Inclined plunging. (*e*) Inclined horizontal. (*f*) Recumbent. (*After M. J. Rickard*, *A classification diagram for fold orientations*, *Geol. Mag.*, *188:23,1971*)

Character of folds. In a section of folded rocks the layers possess different rheological properties; some have apparent stiffness (competency), whereas others behave less stiffly (incompetency). The most competent layers or group of layers control the folding, and the less competent units tend to conform to the fold-form of the most competent units. The least competent units take up the strain between the adjacent fold-controlling units.

The following nomenclature has been devised. The most competent units in the folded sequence are the dominant structural members. Those layers slightly less competent that conform to the fold-form of the dominant structural member are competent conforming members, and the least competent units are passive structural members.

The dominant structural members deform through much of their history of folding so that their thicknesses measured perpendicularly to their boundary surfaces are nearly constant at all points in the fold-form (**Fig. 3*a***). Such folds are said to be concentric. The passive structural members fold so that their thicknesses measured perpendicularly to their boundary surfaces vary from point-to-point in the fold-form, but thicknesses measured in the plane of their axial surfaces and perpendicularly to their axes are relatively constant (Fig. 3*b*). Such folds are called similar. The competent conforming members deform in a manner intermediate between that of concentric and that of similar folds. Under conditions of extreme strain all structural members may be flattened and take on the form of similar folds. Fold-forms when followed either upward or downward in a layered sequence tend to be disharmonic.

The wavelengths of concentric folds are related to the thickness of the dominant structural member. Studies of the relation between wavelength and thickness of the dominant structural member in folds of all sizes conclude $L \simeq 27T$, where L = wavelength and T = thickness of dominant structural member. In a layered sequence, first-cycle folds are controlled by the dominant structural member. Thin conforming competent members have, in addition to first-cycle fold-forms, second-cycle folds (minor folds) developed in the limbs of the first-cycle fold. The wavelengths of these minor folds are related to the thickness of the conforming competent member, that is, the thinner the member the shorter the wavelength.

Strain-stress relationships. Strain in rocks can be measured by comparing the shape of objects of known original form (fossils, pebbles, ooids, reduction spots) to their form after strain. Most studies of

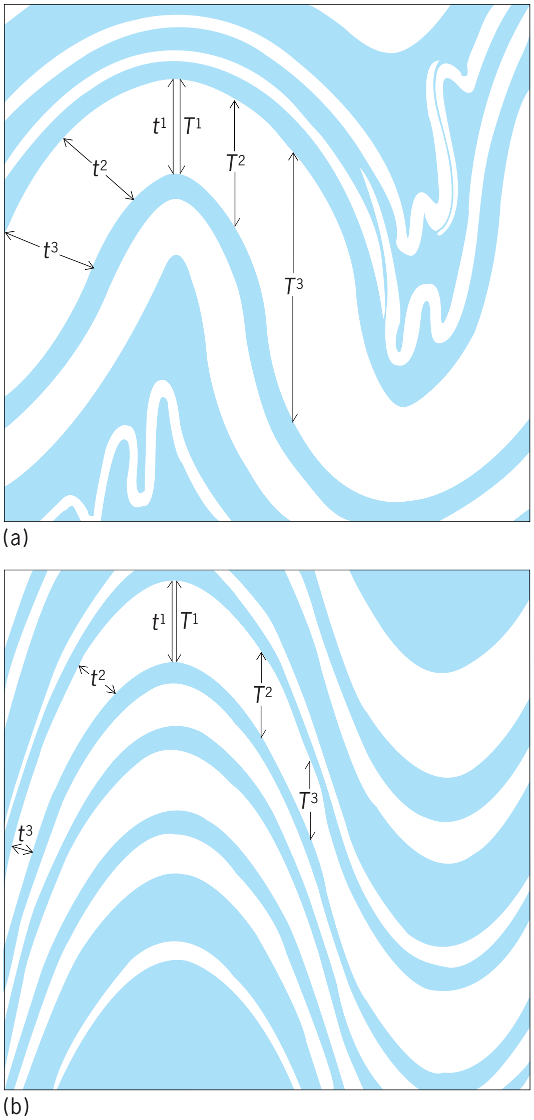


Fig. 3. Geometrical features of (*a*) concentric folds and (*b*) similar folds. Here t^1, t^2, and so on, are thicknesses measured perpendicular to boundary surface of folded layer. T^1, T^2, and so on, are thicknesses measured in axial surface and perpendicularly to fold axis. (*After J. G. Ramsay, The geometry and mechanics of "similar" type folds, J. Geol., 70:310, 1962*)

strain in folds have dealt with rocks that must be considered to be passive structural members in regions of intense folding. Where the folding has not been complicated by more than a single episode of folding, the orientations of the strain axes are as follows: the greatest direction of shortening is perpendicular to the axial surface; the greatest direction of extension lies in the axial surface and is subperpendicular to the axis; and the intermediate strain direction is perpendicular to the other two directions. Strains as great as 60 to 120% are recorded. Doubtless, detailed studies of strain in competent units would have orientations that differ with position within the unit, but such observations are generally beyond the resolution of field studies.

The measured strain orientations in folds of many regions are consistent with folds developed by axial compression of the layered sequence. Theoretical models have been derived on this assumption. Such models also assume the following: the structural members are ideally elastic or viscous; strains are infinitesimal; the thickness of the individual structural members is small relative to the thickness of the layered sequence; the axial load is constant throughout the length of the structural member; vertical normal stresses vary continuously; the fiber strains are linearly related to the distance from the center of the dominant structural member; and the layered sequence is confined by rigid boundaries. The wavelength of the multilayered sequence is given by Eq. (1), where L = wavelength; T = total thickness of multilayered sequence; t = thickness of dominant structural member and associated members; b = breadth of structural members; n = number of cycles of fold; B_{h} = horizontal modulus of elasticity; B_v = vertical modulus of elasticity; and I = moment of inertia. The critical load necessary for buckling with the wavelength given by Eq. (1) is expressed as Eq. (2), where P = critical axial load and G_a = modulus of shear.

$$L = 2\pi \left(\frac{T}{n\pi}\right)^{1/2} \left(\frac{B_b I}{btB_v}\right)^{1/4} \qquad (1)$$

$$P = 2\left(\frac{n\pi}{T}\right)(btB_vB_bI)^{1/2} + G_a tb \qquad (2)$$

Equations for viscous structural members are similar, but time becomes a significant parameter in them, and there is no critical axial load for buckling, although the amount of deflection is sensitive to the magnitude of the axial load. Equations based on both elastic and viscous structural members give wavelengths and numbers of dominant structural members per layered sequence in good agreement with observed characteristics of folded sequences.

Folding may be simulated with computers by programming suitable mathematical equations that relate the necessary boundary conditions, physical properties, and selected effects. These equations can be solved for any set of desired parameters. Two-dimensional maps of the distribution of stress and strain have been constructed for a single dominant structural member embedded in a viscous medium during folding (**Fig. 4**). The strain directions in simulated folds with finite amplitudes and steep limbs have a remarkable similarity to the distribution of strains measured in folded strata. *See* ROCK MECHANICS.

Minor folds and cleavage. Many first-cycle folds are accompanied by higher-cycle folds, called minor folds, and by cleavage. Empirical data indicate that, where these features formed in the same episode of deformation, the axes of minor folds are oriented approximately parallel to the axes of the first-cycle folds. Minor folds also have symmetries that vary with position in the first-cycle folds; they are asymmetrical in the limbs of the larger folds and symmetrical in their hinges.

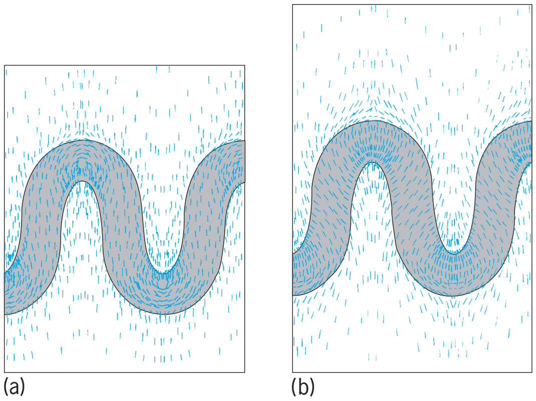


Fig. 4. Computed two-dimensional fold model after 100% shortening. (*a*) Stress relations; short lines are perpendicular to the maximum normal stress. (*b*) Strain relations; short lines are perpendicular to direction of maximum shortening.

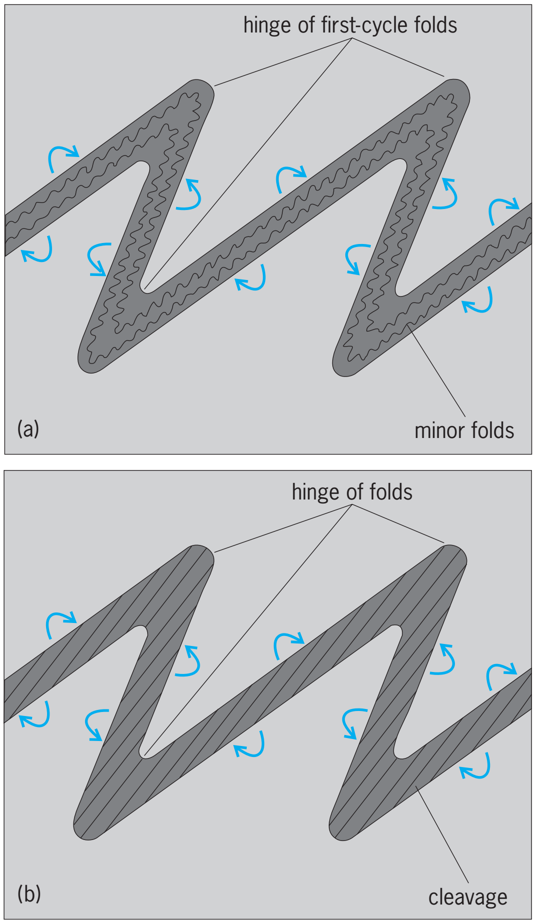


Fig. 5. Relationship between large folds and (*a*) minor folds and (*b*) cleavage; arrows indicate direction of deflection of adjacent hinge relative to limb observed.

The direction of deflection in the hinges of adjacent lower-cycle folds can be inferred from the minor folds (**Fig. 5***a*). The short limbs of the minor folds face the hinge of the larger fold to be considered, and the direction of deflection in this hinge relative to the limb observed is that given by following the long limb into the short limb of the minor folds.

Investigations also show that cleavage, which is manifest by closely spaced surfaces of rupture, approximately parallels the axial surfaces of folds, if these features were formed in the same episode of deformation. The geometric relation between cleavage and bedding is such that their intersection is approximately parallel to the axis of the fold. The direction of deflection in the adjacent hinges of folds can be inferred from the cleavage-bedding relationship (Fig. 5*b*). The acute angle between cleavage and bedding point to the hinge to be considered, and the direction of deflection in this hinge relative to the limb observed is that given by following the trace of bedding into the trace of cleavage around the acute angle between them. Generalizations of the relationships between minor folds and cleavage to larger folds allows the reconstruction of large folds from scattered observations.

Multiple folding. In many regions, layered rocks have had a long history of deformation including several episodes of folding. Because younger deformations do not completely obliterate earlier deformations, a complex array of folds develops. The relative timing of the various folds can be ascertained by their mutual relationships; the axial surfaces of older folds are folded by younger folds, and the axial surfaces of the youngest folds are not deflected (**Fig. 6**). Locally, folds of different ages may be directly superimposed, and it is at these localities that the relative ages are most readily determined. Once the relative timing of folding has been established at one or more localities, the relative ages can be applied to other localities on the basis of fold style. The orientations of superposed folds depend on the orientation of the forces producing them and on the attitude of the surfaces that are folded.

Fold systems. Folds generally do not occur singly but are arranged in festoons in mobile belts with lengths of thousands of miles and widths of hundreds of miles. The folds of these belts, or fold systems, commonly consist of great complex structures composed of many smaller folds. In this regard, a large

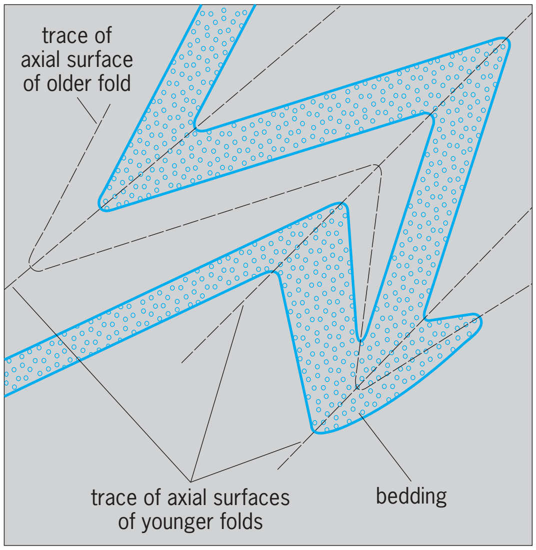


Fig. 6. Superimposed folds showing their relative ages.

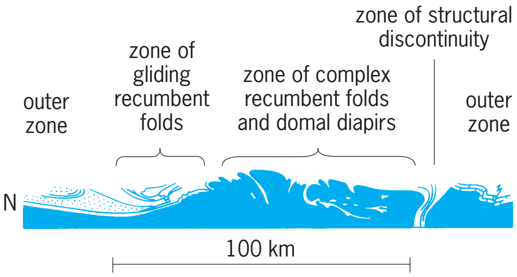


Fig. 7. Section through the Alpine Fold System. 1 km = 0.6 mi. (*After J. I. Dewey and J. M. Bird, Mountain belts and the new global tectonics*, *J. Geophys. Res.*, *75:2639, 1970*)

anticlinal structure consisting of numerous smaller folds is called an anticlinorium, and a large synclinal structure consisting of many smaller folds is termed a synclinorium. The trends of these smaller folds are more or less parallel, although the relationships may be complex. In some regions, folds belonging to several generations may be present, and locally these may have divergent trends.

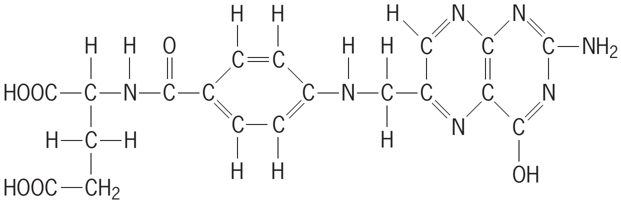
Although comparisons of sections across fold systems are difficult to make because no two fold systems are identical and many are incompletely exposed for study, certain generalities emerge. Most fold systems are asymmetrical with respect to a zone of structural discontinuity (**Fig. 7**). The fold system is more extensively developed on one side of this zone than on the other, and the direction of structural transport on either side is directed toward the margins of the fold system. On the less-developed side of the zone of structural discontinuity, the folds are recumbent or asymmetrical and are accompanied by steeply dipping thrust faults. On the more extensively developed side, and adjacent to the zone of structural discontinuity, is a region of complex recumbent folds and domal massifs. On the outer side of these complex structures, recumbent folds that have moved toward the margin under the influence of gravity are extensively developed, and in the outer margin the folds are asymmetric and inclined and are associated with thrust faults having gentle dips.

The development of fold systems is closely tied to concepts of global tectonics. The favored hypothesis is that of plate tectonics, in which fold systems are formed at converging margins where continents are underriden by oceanic crust, or where collision occurs between a continent and an island arc or between two continents. Alternatively, the hypothesis of an expanding Earth, though less favored, is consistent with much of the geologic and geophysical data. According to this hypothesis, major zones of extension are the loci of gravity-driven diapirs the upper parts of which spread laterally under their own weight to give rise to fold systems. Philip H. Osberg

Bibliography. R. D. Hatcher, *Structural Geology: Principles, Concepts, and Problems*, 2d ed., 1994; D. D. Pollard and R. C. Fletcher, *Fundamentals of Structural Geology*, 2005; J. G. Ramsay and M. I. Huber, *Techniques of Modern Structural Geology*, vol. 2: *Folds and Fractures*, 1987; M. Shankar and G. W. Fisher (eds.), *Structural Geology of Fold and Thrust Belts*, 1992; R. J. Twiss and E. M. Moores, *Structural Geology*, 1992.

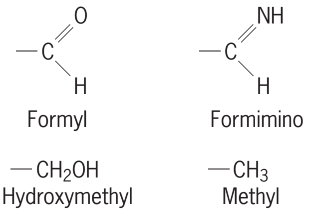
Folic acid

A yellow vitamin, slightly soluble in water, which is usually found in conjugates containing varying numbers of glutamic acid residues. It is also known as pteroylglutamic acid (PGA), and has the structural formula shown below.



It is known that vitamins M and B_c, factors R, S, and U, the *Lactobacillus casei* factor, and a number of others are pteroylglutamates. Folic acid is usually assayed microbiologically by the use of *L. casei* or *Streptococcus faecalis*. Because of the various forms in which the vitamin appears in nature, some having different activity for the two organisms, and because of the difficulty in hydrolyzing the PGA conjugates, it is extremely difficult to determine absolute amounts of the vitamin. *See* BIOASSAY.

Biochemistry. Biologically active forms of folic acid occur as tetrahydrofolic acids. Four of these, N^5-formyltetrahydrofolic acid (the citrovorum factor, folinic acid, or leucovorum), N^{10}-formyltetrahydrofolic acid, N^5,N^{10}-methenyltetrahydrofolic acid, and formiminotetrahydrofolic acid have been found in natural products. Liver and kidney enzymes which catalyze the conversion of folic to tetrahydrofolic acids have been obtained. Ascorbic acid appears to be involved in some of these conversions. Folic acid enzymes are biological carriers of 1-carbon fragments as follows:



These groups are important in the synthesis of purines and pyrimidines, which are components of nucleic acids; in the synthesis of the methyl group of methionine; and in the metabolism of some amino acids, including glycine, serine, histidine, glutamic acid, and phenylalanine.

Nutritional requirements. Folic acid is required by a large number of animals and microorganisms. Many animals obtain enough folic acid from intestinal synthesis so that a dietary source is not necessary. In these species, the feeding of antibiotics or sulfa drugs can precipitate deficiency signs. Folic acid deficiency is usually accompanied by poor growth, anemia and other blood dyscrasias, and gastrointestinal disturbances.

Folic acid is so widespread in nature and intestinal synthesis is so great that a folic acid deficiency in humans because of low dietary intake is probably not very common. With improved diagnostic methods an increasing number of folic acid deficient patients have been discovered. Deficiencies of other nutrients (particularly iron, ascorbic acid, or vitamin B_{12}) leading to impaired folic acid metabolism, difficulties in hydrolysis of folic acid conjugates by intestinal enzymes, or poor absorption may lead to a number of clinical conditions in which folic acid deficiency is involved. These include various nutritional macrocytic anemias, sprue, idiopathic steatorrhea, and pernicious anemia. The daily requirement of humans for folic acid is unknown, but on the basis of food analysis and therapeutic experience, it appears to be about 0.1 mg for the normal adult.

The anemia of pernicious anemia may be prevented by giving a person more than 0.1 mg of folic acid per day, although the neurological manifestations of this disease may progress. Sale without prescription of vitamin preparations recommending more than 0.1 mg folic acid per day has been prohibited by the Food and Drug Administration. *See* VITAMIN.

Folic acid antagonists such as aminopterin, which inhibits the conversion of folic to folinic acid, have been used in the treatment of leukemia and have in many cases brought about temporary remission of the disease. Stanley N. Gershoff

Industrial production. Folic acid is produced commercially by the simultaneous reaction of 1,1,3-tribromacetone with 2,5,6-triamino-4-hydroxypyrimidine and *p*-aminobenzoylglutamic acid. Dibrompropionaldehyde, with the assistance of an oxidizing agent, is sometimes used as a substitute for tribromacetone. Although various by-products are produced in such a reaction, selected conditions of pH and temperature give a reasonably good yield of crude folic acid. Purification of the crude folic acid through the zinc and magnesium salts gives a golden product essentially free of impurities.

Various processes can be used to make folic acid, but this one has been given because of its relative simplicity. The intermediates can be made readily by standard and familiar techniques. Most of the other processes are based on a stepwise conversion of the pyrimidine to the pteridine ring, along with additions to the benzoylglutamic molecule. Folic acid is then produced in a two-component condensation, rather than a three-component condensation. Richard J. Turner

Food

Any processed, partially processed, or unprocessed substance taken into the body primarily for providing nourishment. In addition, factors such as satisfying social needs, achieving psychological ends, and satisfying hunger govern the selection and consumption of foods. Chosen carefully, foods can provide all of the essential nutrients required for the normal functioning of the human body. In this context, food is necessary to provide energy, to provide structural components for building and repairing body tissues, and to regulate body processes. *See* METABOLISM; NUTRITION.

TABLE 1. Functions of essential nutrients

Function	Nutrients involved
Energy	Carbohydrates, fats/lipids, proteins
Structure	Proteins, lipids, minerals
Regulation of body processes	Proteins, minerals, vitamins, water

Nutrient requirements. Essential nutrients, such as carbohydrates, fats, proteins, minerals, vitamins, and water, have specific functions in the human body (**Table 1**). When energy-yielding organic compounds, including carbohydrates, fats/lipids, and proteins, are oxidized in the body, energy is captured in a chemical compound known as adenosine triphosphate (ATP), and then released slowly so that it can be used for respiration, physical activity (work), heat production, and metabolic processes. Enzymes, vitamins, minerals, and water are needed in order for these oxidation reactions to take place. Energy requirements are expressed in terms of kilocalories (kcal), calories (cal), or kilojoules (kJ). Fats and carbohydrates contain only the elements carbon, oxygen, and hydrogen. Since fats contain less oxygen than carbohydrates, they have greater potential for oxidation and thus provide more energy per gram than carbohydrates. Proteins also contain nitrogen, but it does not contribute substantially to the energy value. The physiological fuel value, or the amount of energy generated in the body, of 1 g of protein is 4 kcal (17 kJ); carbohydrate, 4 kcal (17 kJ); fat, 9 kcal (38 kJ). *See* ADENOSINE TRIPHOSPHATE (ATP); ENZYME; VITAMIN.

TABLE 2. Components in each nutrient group for which Dietary Reference Intakes (DRIs) exist

Nutrient group	Essential compounds
Fat	Linoleic acid, α-linolenic acid
Protein	Indispensable amino acids: histidine, leucine, isoleucine, lysine, methionine, phenylalanine, threonine, tryptophan, valine Conditionally indispensable amino acids: arginine, cysteine, glutamine, glycine, proline, tyrosine
Minerals	Calcium, chromium, copper, fluoride, iodine, iron, magnesium, manganese, molybdenum, nickel, phosphorus, potassium, selenium, silicon, sulfate, vanadium, zinc, sodium, chloride
Vitamins	Fat-soluble: A, D, E, K Water-soluble: C (ascorbic acid), thiamin, riboflavin, niacin, B-6, folate, B-12, pantothenic acid, biotin, choline

TABLE 3. Some functions of essential vitamins

Vitamins	Functions
Fat-soluble	
Vitamin A (retinoids, carotenes)	Night vision (preformed vitamin A); antioxidants (carotenoids); growth, reproduction, immune system integrity
Vitamin D (calciferol)	Promotes absorption and use of calcium and phosphorus; rickets prevention
Vitamin E (tocopherols)	Antioxidant, membrane integrity, heme synthesis
Vitamin K	Formation of proteins, particularly for blood clotting
Water-soluble	
Thiamin (B-1)	Coenzyme in energy metabolism, particularly from carbohydrates; antiberiberi factor
Riboflavin (B-2)	Coenzyme in energy metabolism, protein metabolism, redox reactions
Niacin	Coenzyme in energy metabolism, redox reactions; antipellagra factor
Vitamin B-6 (pyridoxine, pyridoxal, pyridoxamine)	Coenzyme for more than 100 enzymes in amino acid and protein metabolism
Folate	Coenzyme in DNA and RNA metabolism; single-carbon transfers, red blood cell formation; reduces risk of neural tube defects in infants.
Pantothenic acid	Coenzyme in fatty acid and energy metabolism
Vitamin B-12 (cobalamin)	Cofactor for methyl (carbon group) transfers; normalred blood cell formation and neurological function
Biotin	Coenzyme in fatty acid and carbohydrate metabolism
Vitamin C (ascorbic acid)	Cofactor for collagen formation; antioxidant; detoxification; scurvy-preventive vitamin
Choline	Needed for structural integrity of cell membranes, lipid and cholesterol transport

Many fats, carbohydrates (glucose and glycogen particularly), and proteins are synthesized in the body. Some of these compounds, as well as minerals and vitamins that are considered dietary essentials, or indispensable nutrients, must be provided by the food supply (**Table 2**). There are also vitamins and minerals that are essential to bodily functions (**Tables 3** and **4**). Dietary Reference Intakes (DRIs), developed by the Institute of Medicine based on scientific data, are reference values for macro- and micronutrients which can be used in dietary planning and assessment. They include four categories: Estimated Average Requirements (EARs), Recommended Dietary Allowances (RDAs), Adequate Intakes (AIs), and Tolerable Upper Intake Levels (Uls). *See* GLUCOSE; GLYCOGEN.

Water is essential in the human diet. Humans can survive for weeks without food, but only a few days without adequate water. Between 50 and 80% of the total body weight of an adult is water. Many foods supply water, but additional fluids must be ingested to meet body needs.

Nutritional guidelines. The Food and Nutrition Board of the Institute of Medicine of the National Academy of Sciences establishes RDAs (Recommended Dietary Allowances) for nutrients; that is, the quantities of certain nutrients that are thought to be adequate to meet the known nutritional needs of practically all healthy persons. To help individuals select adequate diets, food plans have been developed by U.S. Department of Agriculture (USDA) and the Department of Health and Human Services, based on the DRIs for each nutrient, including the RDAs, EARs, and Uls.

TABLE 4. Some of the functions of essential minerals

General function	Elements involved
Acid-base balance	Chloride, phosphorus, sulfate, potassium, sodium, calcium, magnesium
Water balance	Sodium, potassium, chloride
Bone and tooth formation	Calcium, phosphorus, manganese, fluoride
Catalyst for biological reaction and component of essential body compounds, such as hormones	Calcium, magnesium, zinc, copper, iron, molybdenum, cobalt, selenium, iodine, manganes, phosphorus, sulfate
Nerve function	Calcium, sodium, potassium

In 2005, new dietary guidelines, known as MyPyramid, were introduced to help individual consumers make food choices to maintain good health. The guidelines suggest eating a substantial number of servings of cereal-based foods and fruits and vegetables, and smaller amounts of animal products (meats, eggs, dairy) are recommended, compared to earlier published guidelines. The guidelines state that a healthy diet is one that emphasizes fruits, vegetables, whole grains, and fat-free or low-fat milk and milk products; includes lean meats, poultry, fish, beans, eggs, and nuts; and is low in saturated fats, trans fats, cholesterol, salt (sodium), and added sugars. Amounts of food in each group are suggested, together with descriptive information about serving sizes. For example, 2.5 cups of vegetables per day are recommended, particularly the dark-green and orange varieties, such as broccoli and carrots. In addition, balancing physical activity and food intake is emphasized to lower the risk of chronic diseases, such as cancer, diabetes, and cardiovascular disease, which are related to obesity. Reduced fat consumption, particularly saturated and trans fats (see Lipids section below), is also recommended. "Nutrition Facts" labels, on which requirements for terms used on the label to describe nutrient content are specified, aid in the selection of appropriate foods, as required by the Nutrition Labeling and Education Act.

Carbohydrate-rich foods. Plants synthesize carbohydrates by the process of photosynthesis and store them as starches (complex carbohydrates) or simpler sugars (glucose, fructose, or sucrose). The structural components of plants that are nondigestible are called dietary fiber and include cellulose (a glucose polymer that the human body cannot digest), hemicellulose, and pectin. These, together with other complex carbohydrates (such as gums) produced by plants and bacteria, called functional fiber, are classified as total fiber. Functional fiber is beneficial physiologically, as it reduces cholesterol levels. Desirable total fiber intake can be achieved most satisfactorily by consuming fruits, vegetables, legumes, and whole-grain cereals. *See* CELLULOSE; FRUCTOSE; HEMICELLULOSE; SUCROSE.

In most of the world, about 80% of the calories consumed are carbohydrates, primarily from cereal-based foods. (Carbohydrates supply only about half of the calories in the United States.) In addition to being good sources of energy and fiber, whole grains supply B vitamins and minerals and contain small amounts of protein. The most common cereal grains consumed on a worldwide basis are wheat, rice, and corn (maize).

Some grains, such as rice, oats, barley, and buckwheat, are consumed as whole, cracked, or polished grains. Others, such as corn and wheat, undergo a series of mechanical disintegrations that reduce the grain to grits or flour. Whole or cracked grains can be consumed directly as cereals, while grits and flours are baked into bread products, made into pastas, or processed into ready-to-eat cereals. Each processing step adds to the convenience and cost of the final product. Current dietary guidelines emphasize the use of whole grains in breads and cereals. Enrichment of cereal-based foods restores some of the nutrients lost during milling. Thiamin, niacin, riboflavin, folate, and iron are required for a product to be labeled as enriched, while calcium and vitamin D are optional. It is estimated that more than 90% of the flour sold in the United States is enriched. *See* BARLEY; BUCKWHEAT; CARBOHYDRATE; CEREAL; CORN; NIACIN; OATS; RICE; STARCH; THIAMIN; VITAMIN D; WHEAT.

Protein-rich foods. Both animal and vegetable proteins provide the necessary amino acids for the synthesis of proteins in muscle, brain, blood, and other tissues in the human body, as well as hormones and enzymes. Proteins also provide about 15% of the energy in adult diets. Proteins are polymers composed of amino acids. Approximately 20 amino acids commonly are found in food proteins (Table 2), 9 of which are considered indispensable or essential in the diet because they cannot be made in the body. Six are conditionally indispensable; that is, they may be needed in the diet because they cannot be synthesized in sufficient amounts to meet a specific need. The quality of a protein is determined by how well it compares to a protein quality-scoring pattern based on the relative ratio of indispensable amino acids in the food protein and its digestibility, and to a reference protein, such as egg or milk. Legumes and nuts closely resemble meats in their amino-acid balance, but most other plant proteins are limited in at least one of the essential amino acids. The use of complementary plant proteins, such as beans and rice, or supplementation of vegetable proteins with small amounts of high-quality animal or vegetable protein provides adequate amounts of amino acids for human needs. *See* AMINO ACIDS; CHEESE; HORMONE; LEGUME; MILK; PROTEIN.

Lipids. This is the general name for the class of compounds to which fats and oils belong. Fats, chemically known as triglycerides, are composed of fatty acids attached to glycerol. A fatty acid is composed of an even-numbered carbon chain (4–22 carbons), which can be completely bonded with hydrogen (saturated) or can contain some carbon-carbon double bounds (unsaturated; that is, the adjacent carbons each lack one hydrogen). The position of the hydrogen atoms surrounding the double bonds is either *cis* (hydrogen atoms on one side of the bond) or *trans* (hydrogen atoms on opposite sides). The fatty acids of triglycerides in fats or oils are characteristic of the source of the material. All foods contain some lipids, even fruits and vegetables. In fats from animal sources, the majority of the fatty acids are saturated (for example, stearic and palmitic acid), but in vegetable fats, the fatty acids may be mono-, di-, or polyunsaturated. Cholesterol, an important component of cell membranes and precursor for steroid hormones and bile acids, is present only in foods from animal sources. *See* CHOLESTEROL; GLYCEROL; LIPID; TRIGLYCERIDE.

Saturated fats, such as butter, lard, and tallow, are solids at room temperature; polyunsaturated fats (oils) are liquids at room temperature. Saturated fats generally have low smoke points (the temperature at which the fat begins to break down) and are not as useful for frying; oils have higher smoke points. The most common vegetable oils in the United States are extracted from soybeans, cottonseed, corn, sunflower, canola, and safflower, which all have a high proportion of polyunsaturated fatty acids. Hydrogenation of oils for margarine and shortening results in formation of trans fats, which resemble saturated fats in their structure and physical properties. Labeling of trans fat content of foods was required as of January 2006, because of its connection with cardiovascular diseases. Dairy and meats naturally contain some trans fats. More than half of the dietary fat in the United States comes from animal sources; the remainder is from vegetable oils. Slightly less than 40% of the calories in the American diet are from fats and oils. *See* BUTTER; COTTON; RAPE; SAFFLOWER; SOYBEAN; SUNFLOWER.

Vegetable oils can be partially or fully hydrogenated to make solid or plastic fats. In making margarine, some liquid oil is added to hydrogenated fat, as well as color, flavoring, and liquid (usually milk) to emulate the consistency, flavor, and cooking characteristics of butter. Spreads containing lower amounts of fat than butter, or solid margarines, can be made by blending saturated (palm), polyunsaturated (soybean, canola, cottonseed, sunflower, or safflower) or monounsaturated (olive) fats and oils. By using these alternatives, cooking properties may be changed. Properties of hydrogenated shortenings are altered by fat substitution and processes that minimize trans fats in the finished products, so food companies are developing patented technologies to achieve the desired functional properties. Reduced-calorie spreads can also be made by diluting the fat with water or incorporating air to make a whipped product. For baking and frying, shortening can be made from hydrogenated vegetable oils alone or can be blended with animal fat. Superglycerinated shortenings have mono- and diglycerides added to improve baking and emulsifying properties. *See* FAT AND OIL; MARGARINE.

Vegetables and fruits. Recent nutritional guidelines emphasize the consumption of a variety of fruits and vegetables, particularly dark-green and deep-yellow ones. These foods constitute major sources of the carotenoids (such as vitamin-A precursors), ascorbic acid, and many minerals, as well as digestible carbohydrates and dietary fiber. Fruits and vegetables also contain phytochemicals or phytonutrients, which provide a health benefit beyond those of essential nutrients. Vegetables include a variety of edible parts of plants (leaves, stems, flowers, roots, tubers, pods, and seeds) and are usually eaten cooked or raw, as a part of a meal. Fruits of certain plants (tomatoes, eggplant, squash, and peppers) are grouped with the vegetables. Fruits generally contain more sugar than vegetables. Both fruits and vegetables are very good sources of dietary fiber.

Supplies of fresh fruits and vegetables once were limited by their seasonal nature, but the development and refinement of transportation, packaging, and canning and freezing technologies have extended their availability. Freezing or canning of fruits and vegetables is done as soon as possible after harvesting, usually within hours to help preserve the optimum quality and nutritional value. The nutritional value of canned and frozen fruits can be similar to that of "retail market" fresh fruits and vegetables. Modified and controlled-atmosphere storage extends the shelf life of fresh produce. *See* FRUIT.

Foods for health. Although it is unlikely that any individual food or combination of foods can provide complete disease protection, the risk of chronic disease can be reduced by increased consumption of plant-based foods and decreased consumption of fats. Foods that provide essential nutrients, beyond quantities needed for normal growth, maintenance, and development, or contain other biologically active components are called functional foods. These include fruits, vegetables, and whole grains. Because obesity in the United States has increased significantly over the last 30 years, health, nutrition, and food-science professionals are emphasizing healthy eating and exercise to decrease the risk of obesity-related diseases, such as type-2 diabetes and cardiovascular disease.

The food industry is focusing its product development efforts to meet recommendations for weight loss and health maintenance. Formulated foods that are described as fat-free, low in fat, or reduced-calorie depend on the use of fat substitutes and mimetics, as well as high-intensity sweeteners. High-intensity artificial sweeteners, such as acesulfame-K, aspartame, saccharin, and sucralose are used in tabletop sweeteners, beverages, candies, dessert mixes, and dairy products. When these sweeteners replace sugar in food products, bulking agents (such as sorbitol, polydextrose, or cellulose) must be added. Fat substitutes and mimetics contribute fewer calories than fat because they are partially absorbed or nonabsorbed in the intestinal tract. They provide sensory characteristics in foods similar to those imparted by fats. Carbohydrate-based fat substitutes, such as cellulose, gums, dextrins, modified starches, and polydextrose, can be found in a variety of foods, including dairy products, frozen desserts, and salad dressings. Microparticulated protein is made from egg white or milk proteins. The spherical particles produced by the processes are so small that the tongue perceives them as a fluid, with a smooth, creamy texture. Microparticulated protein cannot be used in cooked foods because it forms a gel when heated. Sucrose polyesters are composed of six to eight fatty acids bound to a sucrose molecule. These fat substitutes are heat stable and have the potential for use in baked and fried foods. According to the American Dietetic Association, fat-modified foods can be used as part of a well-balanced, nutritious diet to reduce fat intake. *See* FOOD ENGINEERING; FOOD MANUFACTURING; FOOD MICROBIOLOGY; FOOD SCIENCE.

Barbara P. Klein

Bibliography. R. L. Duyff, *American Dietetic Association Complete Food and Nutrition Guide*, 2d ed., Wiley, New York, 2002; Food and Nutrition Board, Institute of Medicine, National Academy of Sciences, *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids* (*Macronutrients*), National Academies Press, 2005; L. K. Mahan and S. Escott-Stump (eds.), *Krause's Food, Nutrition & Diet Therapy*, 10th ed., Saunders, Philadelphia, 2004; H. McGee, *On Food and Cooking: The Science and Lore of the Kitchen*, Scribner, New York, 2004; N. N. Potter and J. Hotchkiss, *Food Science*, 5th ed., Springer, 1999; G. M. Wardlaw, *Contemporary Nutrition: Issues & Insights*, 5th ed., McGraw-Hill, Boston, 2003.

Food allergy

An "altered reaction" to a food substance to which people typically do not react. Although the term "allergy" is sometimes used loosely for many different kinds of response, strictly speaking it refers to an immunological response mediated by the IgE class of immunoglobulins, or antibodies.

As many as 8% of children and 2% of adults in the United State have had food allergy symptoms. Each year about 30,000 people visit emergency departments, and an estimated 150 die from allergic reactions. Allergy to food is similar to that to other environmental allergens such as pollens, but there are additional specific risks, symptoms, and immunological mechanisms.

IgE antibodies and allergic mechanisms. IgE immunity plays an important role in protection against intestinal pathogens, primarily worms. Although everyone produces some IgE antibody to most environmental antigens, especially those that enter the body through the mucous membranes of the respiratory and gastrointestinal tracts, people with allergies make more. This process is thought to be driven by excessive activity of helper T cells (Th) of the type 2 class (Th2) compared with Th1 cells for reasons that are both genetic and environmental. Although about 15% of the general population has some form of allergy, if both parents of an individual are allergic the individual's risk of having allergies increases to 50%. Allergies are more common in rich than poor countries and in urban than rural settings. This evidence supports the "hygiene hypothesis," the idea that exposure to environmental antigens is important in the first years of life to stimulate the Th cells and achieve an optimal balance between Th1 and Th2 cell activities. *See* ANTIBODY; CELLULAR IMMUNOLOGY.

There is very little IgE free in the blood, because as soon as it is made it binds to surface receptors on mast cells, connective tissues cells containing granules of biologically active chemicals (mediators). Thus in a person with allergies, many mast cells will be coated with IgE against one or a few antigens which, upon entering the body, will cross-link adjacent IgE molecules. This signals the mast cell to immediately release the contents of its granules, which include large amounts of histamine. Within a few hours, the mast cells also generate proinflammatory mediators, such as cytokines and leukotrienes. In response to histamine and other mediators, small blood vessels dilate and become somewhat leaky, causing local tissues to swell with fluid; bronchioles constrict; and gastrointestinal smooth muscle contracts. Histamine also causes itching by stimulating nerve endings. *See* ALLERGY; CYTOKINE; HISTAMINE; INFLAMMATION.

Triggers and symptoms. Only eight food types trigger 90% of allergies. Peanuts, tree nuts, fish, and shellfish are the most common allergens in adults. Children can be allergic to these foods, as well as to milk, eggs, soy, and wheat. For reasons that are not understood but may have to do with the maturation of Th2 cells, children frequently outgrow their allergies as they enter adolescence. Adults rarely do, though severity of symptoms may wax and wane.

The symptoms of food allergy are most common in the gastrointestinal tract but can also be seen on the skin, in the lungs, and systemically (throughout the body). They almost always occur soon (within seconds to 2 hours) after eating the trigger food. Gastrointestinal symptoms range from mild discomfort to cramping, nausea and vomiting, and diarrhea. Respiratory symptoms include coughing; wheezing; and a congested, itchy, or runny nose. The skin may display urticaria (or hives), localized itchy swellings surrounded by a flare of dilated capillaries. Occasionally, large hives may develop beneath the surface of the skin, a condition known as angioedema. The most feared symptom of allergy is anaphylaxis, a systemic allergic reaction that occurs when enough histamine is released from mast cells all over the body that generalized blood vessel dilation results in a drop in blood pressure. The patient may experience just some light-headedness or actually may go into anaphylactic shock, lose consciousness, and require emergency resuscitation. *See* ANAPHYLAXIS.

Oral allergy syndrome. A recently recognized special case of food allergy is oral allergy syndrome, an allergic reaction to certain proteins found in a variety of fruits, vegetables, and nuts that affects people with sensitivity to pollen. The reaction occurs almost immediately on putting the offending food in the mouth. The antigens seem to be able to pass through the mucous membranes of the mouth and gain rapid access to local mast cells, and eventually to mast cells at more distant sites. In most cases, these particular antigens are destroyed in the stomach; thus individuals experience only oral symptoms, including tingling of the lips and tongue, and itching and sometimes swelling of the lips. Though usually just a nuisance, oral allergy syndrome may lead to more severe and general symptoms and should not be ignored by either patients or their health care providers.

Studies have repeatedly shown that the foods that cause oral allergy syndrome contain proteins that are structurally similar to proteins in certain pollens. It is thought that the patient becomes sensitized (that is, makes IgE) in response to the pollen, which may or may not in itself trigger allergic symptoms. The patient's immune system has difficulty distinguishing between the pollen proteins and the food proteins, and triggers an allergic reaction in response to the food proteins. This is known as a cross-reaction, as may occur between inhaled birch pollen and ingested apples, nuts, and carrots; between grass pollen and green beans and peas; or between ragweed pollen and cantaloupe, honey dew melon, banana, or chamomile. The food reactions are often worse in pollen season.

Diagnosis. Allergists work to discover which food in a varied diet is triggering symptoms. A careful medical history of the patient is taken, and suspicious foods are temporarily eliminated. One by one they can be reintroduced and the patient observed. This is not done when the reaction has been severe; in these cases, skin tests may be performed (and only by specialists in the case of very severe symptoms). A skin test involves application of the antigen to the skin by injection or scratching. The development of a hive after about 15 minutes is considered a positive result. Although this does not prove the symptoms are from that antigen, it raises the index of suspicion. For most common food allergens, a completely safe blood test known as the radioallergosorbent test, or RAST, is available. The gold standard is the placebo-controlled double-blind test, in which food is placed in unmarked capsules and ingested while the patient is under careful observation in an experienced hospital or clinic setting.

Food intolerance. Food allergy is commonly misdiagnosed, often by the patient who has had some sort of unpleasant reaction to a food. Food intolerance is the generic term for such reactions, which are not mediated by IgE, and it is perhaps five times more common than true allergy. Certain food such as fish in the tuna and mackerel family, if not properly refrigerated, can develop high enough levels of histamine to cause symptoms indistinguishable from food allergy. Some people are intolerant to lactose, a sugar found in milk, and experience bloating and diarrhea after eating dairy foods because they are deficient in the enzyme lactase, which is necessary for lactose digestion. In some people, monosodium glutamate, a flavor enhancer, causes flushing of the skin and headache; and sulfites, which are used as preservatives, can trigger asthma. Gluten intolerance is an immunological reaction to the gluten protein in wheat, though the reaction is not mediated by IgE. It is not uncommon to develop a "conditioned aversion" to a food that once made us ill, and some people experience symptoms if the food is even mentioned.

Treatment and prevention. Treatment depends on whether the problem is acute and life-threatening, or mild and chronic.

In emergency situations, in which an individual begins to experience shock and labored breathing, drugs, such as epinephrine, which constrict blood vessels and dilate bronchioles, are essential. Fluid replacement and the administration of antihistamines, oxygen, and glucocorticoids may all become necessary. Once the patient is stable, efforts must be made to identify the offending food because strict avoidance is the best preventive measure. Since people with severe allergies (such as to peanuts and shellfish) can have an anaphylactic reaction to a skin test with the trigger substances, expertise with these procedures is mandatory. Specialized centers can perform sensitivity tests in a controlled and safe manner. It is advisable for people with a history of anaphylactic reactions to wear medical alert bracelets or other identification and carry preloaded syringes containing epinephrine at all times.

For milder conditions, avoidance of the offending food is also the mainstay of treatment. This is preferable to continuous use of antihistamines, as is recommended for people with inhalant allergies. It is important for health care providers to make people with food allergy aware, without inducing panic, that mild reactions do rarely progress to severe or even anaphylactic intensity. For example, angioedema is an early warning sign of trouble, as the large hive-like swellings of the face can sometimes migrate to the throat, blocking the airway. If symptoms are progressing, it is vital to seek medical help at once.

Most children will outgrow their food allergies. Adults do not seem to do so, although severity of symptoms varies. For example, a person who can eat a few shrimp without trouble may develop diarrhea and wheezing after eating large quantities. Most people learn what works for them and can usually manage a reasonable diet without much extra vigilance. Those with severe allergy, for example to peanuts, must be very careful, read all labels, ask about ingredients, and nevertheless always be prepared for an allergic reaction.

Recently a new monoclonal antibody against IgE, called omalizumab, has been approved for treatment of severe asthma, and is being evaluated for use by peanut-sensitive patients. In early trials, patients were able to tolerate controlled doses of peanut that would otherwise have triggered anaphylaxis. *See* MONOCLONAL ANTIBODIES. J. John Cohen

Bibliography. Articles about all aspects of food allergy, *Curr. Opin. Allergy Clin. Immunol.*, vol. 5(3), Lippincott Williams & Wilkins, June 2005; J. Brostoff and L. Gamlin, *Food Allergies and Food Intolerance: The Complete Guide to Their Identification and Treatment*, Healing Arts Press, 2000; D. Y. Leung et al., Effect of anti-IgE therapy in patients with peanut allergy, *N. Eng. J. Med.*, 348:986–993, 2003; U.S. Department of Health and Human Services, National Institute of Allergy and Infectious Diseases, *Food Allergy: An Overview*, NIH Publ. no. 04-5518, July 2004.

Food engineering

The application of engineering concepts and principles to the conversion of raw foods into safe consumer products of the highest possible quality. The entire spectrum of food engineering is associated with operation and maintenance of food processing plants as well as sophisticated research involving process design.

The applications of engineering in food handling, processing, packaging, and distribution can be described in terms of unit operations. There are many different unit operations associated with the conversion of raw food materials to consumer products. The movement of foods and other materials within the processing plant requires the use of unique equipment and processes. For example, special sanitary pumps are used to transport liquid foods, and the material-handling equipment for solid foods requires careful design for product-contact surfaces.

The importance of thermal treatments for food preservation requires that a broad range of heat-exchange equipment be used. Heat exchangers for liquids are unique in terms of sanitary design and cleanability of surfaces following thermal processing. A special component of thermal preservation is the design of thermal processes.

Several unit operations involve heat transfer in order to achieve the desired preservation even though storage stability is not the direct result of thermal treatment. An excellent example is the freezing process, where removal of thermal energy reduces product temperatures to levels where deterioration reactions are significantly inhibited. Concentration processes achieve a degree of preservation by reducing the availability of water for deterioration reactions, although the primary aim is reduction of liquid-product mass and volume. Although traditional concentration processes have used thermal energy to evaporate water, membranes of various types are now used to achieve the same results. The preservation of food products is achieved by reduction of the water content to low levels by means of dehydration processes which use thermal energy. These processes are applied to liquid foods and to products that are naturally solid. *See* FOOD PRESERVATION.

Another series of unit operations is used to alter the product composition or structure in some manner. These include separation, mixing, and extrusion. Separation processes are designed to divide food products into two or more components. While a variety of physical or chemical properties of the product components are used in the various separation processes, two of the most important processes are filtration and extraction. Filtration, a physical process, has several applications in addition to its use for separating product components. Extraction is most often designed to remove a specific or unique product component for use in a separate operation or product formulation. After separation, the final product is obtained through the use of a mixing process which includes a variety of equipment types. Finally, the extrusion process involves the use of both thermal and flow properties to achieve product preservation as well as some specified set of structural and textural characteristics.

The importance of cleaning and sanitation must be emphasized due to direct relationships to final product quality. The required operations vary considerably depending on the type of product handled and the type of equipment used. The processes required to manage the wastes generated during food handling, processing, packaging, and distribution are all similar, and many of the waste-handling and treatment operations are the same as those used directly with the food products.

The final operation to which the product is subjected before distribution is packaging. The package barrier is important for maintaining food products at desirable quality levels. Food packaging involves the selection of equipment needed for placing the product in the package as well as the selection of packaging material needed to protect the product in an optimum manner.

An engineering input to food handling, processing, packaging, and distribution that is applied to almost all unit operations is process control. The use of instrumentation and associated electronic controls has a significant impact on the efficiency of all components of the food delivery system. *See* UNIT OPERATIONS.

Dennis R. Heldman

Pumps and Pumping

Pumps are one of the most common pieces of equipment in food processing facilities. Fluid foods with widely differing characteristics must be pumped between many different unit operations before a final product results. Some pumps are selected for a single application or duty, and others are used for various products and flow rates.

Sanitary pumps. Pumps used in handling food must be sanitary, that is, specifically designed to handle food or related biological materials, such as pharmaceutical products. Sanitary pumps are designed for easy and frequent cleaning. They are easily dismantled by removing wing nuts by hand or by using simple tools. In addition, the units are mounted to pipelines with special clamps and disconnects designed for rapid breakdown. Pumps may also be designed for cleaned-in-place systems, which reduce the need for frequent dismantling. These self-cleaning units are constructed with grooves which route sanitizers and create turbulence to help dislodge solids.

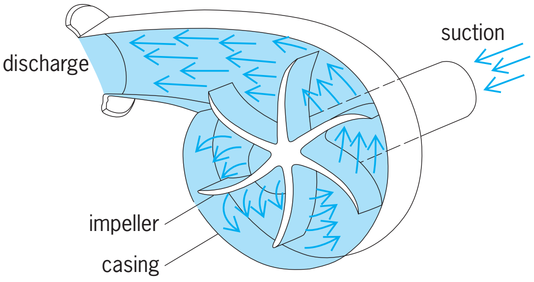


Fig. 1. Internal view of a centrifugal pump. (*After C. R. Westaway and A. W. Lommis, eds., Cameron Hydraulic Data, 16th ed., Ingersoll-Rand, 1984*)

Sanitary pumps must be manufactured from materials that are resistant to corrosion and the chemicals used for cleaning and sanitation. Pump rotors and casings are commonly available in stainless steel; rotors are also available in natural or synthetic rubber. Sanitary pumps should have no internal wear, and the lubrication systems for bearings and other moving parts should be completely sealed to prevent contact with the food being pumped. This is typically accomplished with single or double O-ring–type seals which may incorporate flushing systems for vacuum applications or when pumping hot fluids.

Centrifugal pumps. Pumps used to handle fluid foods will generally fall into one of the two classes: centrifugal or positive displacement. Centrifugal pumps differ widely in size and internal design; however, they all have the same basic principle of operation. The pump (**Fig. 1**) consists of an impeller rotating within a casing and attached to a shaft that is driven externally by an electric motor. Liquid enters the pump at the center of the rotating impeller and is accelerated to a high velocity by the movement of the impeller vanes. The resulting movement—from the center to the outer edge of the impeller—greatly increases the kinetic energy of the fluid. The fluid velocity is reduced when the material is discharged by the vanes to the casing, which causes some of the fluid kinetic energy to be converted to pressure. This increased pressure causes fluid movement from the discharge section of the pump.

Liquid is delivered from a centrifugal pump in a steady stream, that is, without pulsations. Wear is minimized because the pump has no internal rubbing parts and clearances are relatively large. Centrifugal pumps are not self-priming, and develop a limited pressure when operated at a constant speed.

Centrifugal pumps are very versatile and can handle all liquids except highly viscous ones. Large clearances make them useful in pumping liquids containing solids such as cherries, oysters, whole fish, beans, and olives. Overagitation of liquids may be a problem with centrifugal pumps if operated against high pressure. In addition, centrifugal pumps may incorporate

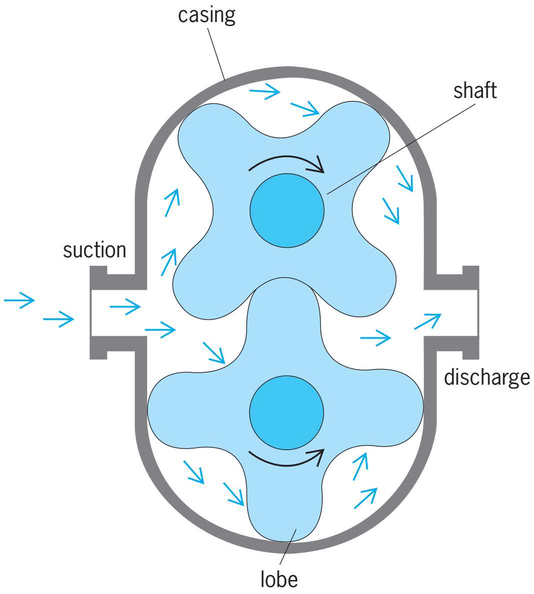


Fig. 2. Internal view of a four-lobe rotary pump.

rotor designs that provide mixing, emulsification, or dispersion as part of the pumping operation. *See* CENTRIFUGAL PUMP.

Positive-displacement pumps. Positive-displacement pumps operate with close internal clearances, and commonly use revolving gears, vanes, lobes, or screws in a fixed casing to bring about fluid movement (**Fig. 2**). In contrast to centrifugal pumps, fluid may be delivered in a pulsating flow. Lobes are attached to shafts that rotate in opposite directions. During rotation, fluid is trapped between the lobe and the casing at the suction end of the pump, then moved (or displaced) to the discharge end where volumetric displacement by the lobes causes the fluid to exit the system. A given quantity of fluid is delivered with each rotation of the shaft. Rotary pumps are self-priming and capable of producing very high pressures; however, excessive pressures can occur if the pumps are operated against a closed valve. These units are very useful in handling highly viscous fluids, and are also used for fluids with low viscosity. Due to close working clearances, rotary pumps are limited to fluids that do not contain large particles.

Rotary pumps, available for aseptic processing systems (that is, for sterile products placed in sterile containers), incorporate design features which eliminate the possibility of undesirable microorganisms entering the product through the pump. These units operate by having steam or a sterile solution circulate between double seals at the ports, cover, and shaft of the pump.

Positive-displacement pumps may combine self-priming features with the transfer capabilities of centrifugal pumps. This may be accomplished by using a flexible neoprene rotor with a pump casing having a noncircular cross section (**Fig. 3**). Rotary action of the impeller causes material to be picked up at the inlet port and transferred in discrete units to the outlet. When the rotor contacts the offset section of the casing, the rotor bends, causing the food to be squeezed from the pump. This type of design tends to minimize damage to suspended solids.

Positive-displacement pumps may be designed with a progressing cavity. One commercial design (**Fig. 4**) involves a screwlike rotor that turns within a double-thread helical stator to form a progressing cavity in the pump. Pumps based on this principle tend to be very effective in handling highly viscous, abrasive, or shear-sensitive fluids. A second design, the twin piston pump (**Fig. 5**), alternates pumping from one piston to the other. It has large, smooth passages that allow it to pump highly viscous liquids, fruit and vegetable products with large particulates, dough, and meat products. Moving sleeves and large ports minimize product damage. *See* DISPLACEMENT PUMP.

Pumping. Pumps are driven by electric motors and provide an energy input to the fluid-handling system. Power requirements for pumping are calculated by using the mechanical-energy-balance form of Bernoulli's equation. Energy inputs must increase if the potential, kinetic, or pressure energy of the fluid is increased. In addition, the pump must overcome energy losses due to friction caused by flow through straight pipe, valves, fittings, and processing equipment, such as filters and heat exchangers. Pumping fluid foods must often be given special consideration because many products exhibit non-newtonian behavior. Common problems include time dependency and the presence of a yield stress. These properties are taken into account when calculating pipeline energy losses due to friction and kinetic energy differences. In addition, many materials (such as an emulsion like mayonnaise) are sensitive to shear, which may limit engineering design options, including pump speed and pipe diameter. Product data such

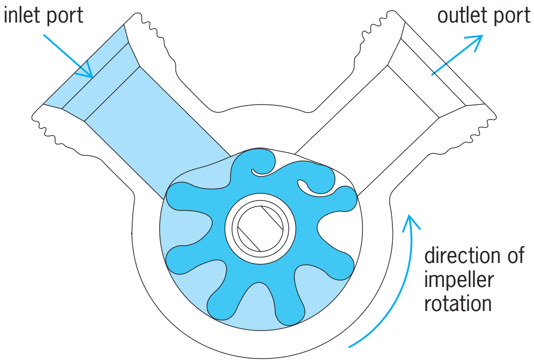


Fig. 3. Rotary pump with a flexible impeller and off-set casing.

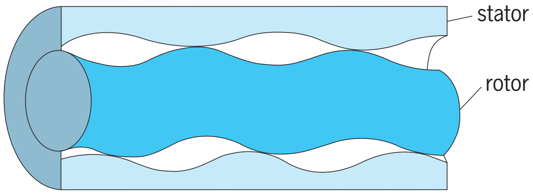


Fig. 4. Rotor and stator for pumps operating to form a progressing cavity.

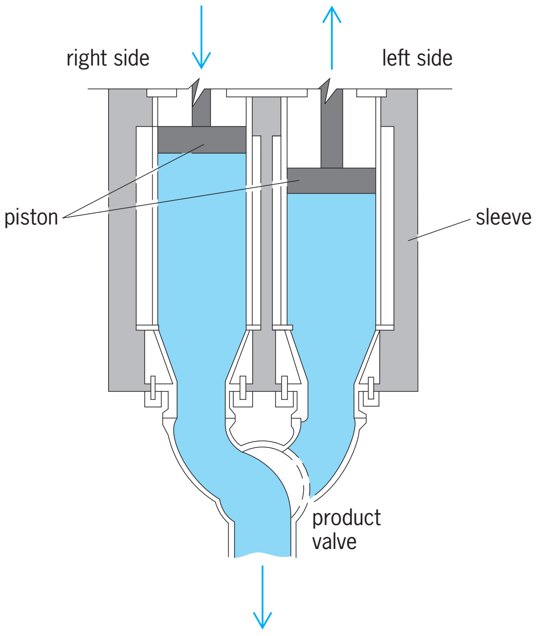


Fig. 5. Twin-piston pump. (*Marlen Research Corp.*)

as temperature, particle size, specific gravity, corrosiveness, and abrasiveness must also be considered in pump selection. *See* BERNOULLI'S THEOREM; FLUID-FLOW PRINCIPLES; FLUIDS; NON-NEWTONIAN FLUID; PUMP; PUMPING MACHINERY.

James F. Steffe

Material Handling

Material handling is a term used to describe the movement of materials through various stages of processing. It includes movement of raw material from a supply location to a processing plant; movement of material through different stages of processing within a plant; and movement of finished product to a warehouse or to storage. In material handling it is imperative that the movement of a known quantity of material to a preselected location occurs at the desired time while maximizing the economy of space.

Equipment used for the mechanical handling of materials may be classified as conveyors and elevators, hoists and cranes, trucks, and pneumatic handling equipment.

Fig. 6. Belt conveyor. (*Aseeco Corp.*)

Conveyors and elevators. Conveyors and elevators carry material in a limited but continuous stream. Belt conveyors (**Fig. 6**) involve endless belts running about end pulleys and supported along the length by a series of securely mounted pulleys called idlers. Commonly the belts are made of rubber, neoprene, canvas, or stainless steel alloys. In situations where continuous cleaning is necessary, the return side of the belt may be scrubbed with roller brushes and washed with water sprays. A variety of modifications allows conveyors to be customized for special requirements. Belt conveyors with stainless steel cleats attached to rubber belts are used to move dry products. A slat-type belt may be useful to convey chunky products, such as cut-up chicken. Conveyors may also be mounted on scales for continuous weight measurements.

Screw conveyors use a helical screw that revolves and pushes the product through a trough or a box. Commonly, these conveyors consist of stainless steel ribbons wound helically around the central shaft or pipe. These conveyors may be mounted horizontally at an inclination or even vertically. They are useful in conveying dry and slightly moist products, for example, flour, cereals, and comminuted meat.

Bucket elevators consist of steel, malleable iron, or synthetic plastic buckets attached to endless chains or belts running about end sprockets or pulleys (**Fig. 7**). Usually the buckets are attached at equal spacings. Bucket conveyors raise material in a vertical plane and are used to transport dry products, such as coffee beans, sugar, and salt.

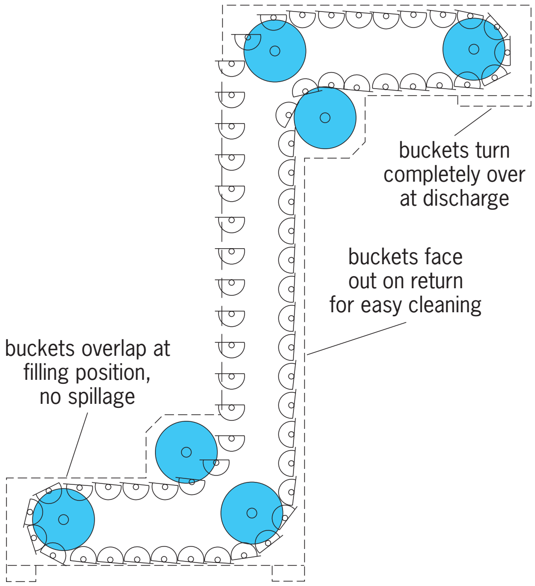


Fig. 7. Schematic of a bucket conveyor. (*Aseeco Corp.*)

To convey delicate products that must be handled gently to avoid breakage or product segregation, vibratory conveyors are used (**Fig. 8**). Vibratory conveyors employ the principle of natural frequency to move the product gently and eliminate any possibility of particle separation. The conveyor pans are supported by springs, and the drive mechanism is tuned to resonance with the natural frequency of both springs and pans. The product moves in a "throw-and-catch" mode. During the upstroke of the pan, the product is lifted gently and pitched forward by the pan. In the following downstroke, the product is caught by the pan before it reaches the bottom of its stroke. This procedure helps in avoiding impact of the product with the pan and thus minimizes physical damage. In addition, the wear of the conveyor itself is small when it is used to convey abrasive materials. These conveyors tend to be self-cleaning and avoid buildup of material. In handling wet products, the self-cleaning feature is particularly attractive as it reduces shutdown time for cleaning.

The application of computers to control conveyor movement has become increasingly popular in plants that handle glass bottles at high speeds. For example, a programmed computer may be used to control the number of bottles on each conveyor section, and to control the speed of the conveyor belt and any associated processing machinery.

Fig. 8. Vibratory conveyor. (*Aseeco Corp.*)

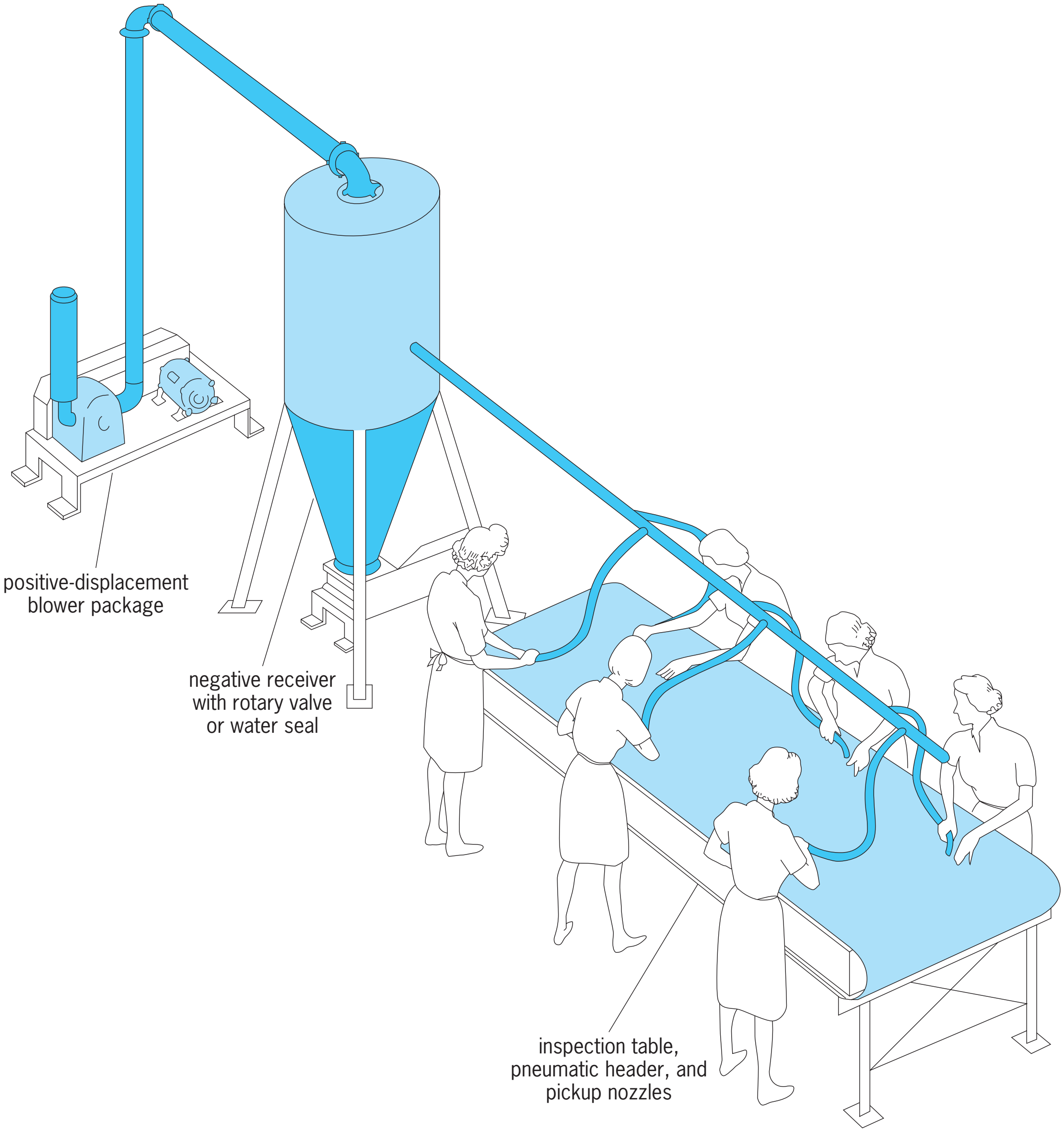


Fig. 9. Pneumatic conveying system. (*Temco, Inc.*)

Computerized control allows optimal line efficiency, lower noise level, reduced breakage by gentler handling, reduced power consumption, and reduced manual work for operators. *See* BULK-HANDLING MACHINES; CONVEYOR; ELEVATING MACHINES.

Hoisting. Hoisting towers are primarily used for unloading purposes; the hoisted material is transported according to whether the tower is of the stationary or traveling type. Simple mast rigs are used in loading and unloading loose material with buckets or scoops. Electrically operated overhead cranes are used to raise or lower heavy loads and distribute them with the trolley traveling on a bridge; the bridge itself moves on the supporting rails. *See* HOISTING MACHINES.

Industrial trucks. These are used in handling food at all stages of production. In the food industry, the term truck is used to include a large range of vehicles—from manually operated flat-pallet trucks, such as the ones used in supermarkets, to battery-powered equipment used in stacking loads of almost 3 tons (2.7 metric tons) to heights up to 30 ft (10 m) in warehouses. Most applications inside a food warehouse or processing facility preclude the use of an internal combustion engine as a power source, because it is mandatory to avoid product contamination from engine exhaust gases. Battery-powered equipment is used mostly indoors, while diesel-powered trucks are used outside. In comparison to a truck with an internal combustion engine, an electric truck is relatively slow and requires battery recharges between shifts. The trucks themselves must be protected from any hostile environment, such as contact with chemicals that may be used in the processing plant. For example, in a fish processing plant, brine can be hazardous to the truck body. Special construction features are used to reduce such damage. *See* INDUSTRIAL TRUCKS.

Pneumatic systems. These systems can be broadly classified as vacuum systems (negative-air systems) and pressurized systems (positive-displacement systems). The selection of a system depends on the product characteristics and rate of conveyance. In a negative-pressure system, vacuum is created by operating a positive-displacement blower located downstream. A rotary air-lock valve connected to the blower is also used to discharge the product. This type of system has been successfully used for removing wastes from processing plants (**Fig. 9**). Positive-airflow systems use a blower connected to a rotary air-lock valve at the feed end. The product is blown into the conveyor pipe. At the discharge end a cyclone is used to separate product from air; the product is allowed to fall by gravity into bulk bins. These systems are popular in conveying cereal grains and individually quick-frozen fruits and vegetables.

Tankers. In transporting granular and powdered products, pressurized road tankers have become more widely used. These tankers avoid the time wasted on filling bags at the warehouse and later emptying bags at the customer's location. Instead the product is loaded pneumatically into a tanker, which is then weighed and driven to its destination (the customer) where it is pneumatically emptied. Pneumatic transport offers many advantages, including reduced packing and unpacking costs, elimination of industrial injuries due to lifting of bags, less handling of product, and reduced rodent infestation. *See* MATERIALS-HANDLING EQUIPMENT. R. Paul Singh

Hydrocyclones

Hydrocyclones (also known as liquid cyclones) are stationary devices that utilize fluid dynamic forces to separate solid from liquid or two solids of differing density suspended in a liquid. The most effective use of hydrocyclones is in the separation of protein in a water suspension of starch in starch manufacturing.

In wet milling of corn, primary separation of starch and protein is performed with large disk-nozzle centrifuges. The crude starch fraction (1.5–3% protein) is then countercurrently washed (**Fig. 10**) by using small (0.4–0.6 in. or 10–15-mm diameter) hydrocyclones to remove most protein and other soluble materials (solubles). The smaller less-dense protein conglomerates remain in the overflow, while the heavier starch granules settle and exit the system in the underflow. *See* CENTRIFUGATION.

In the starch washing system, clean wash water enters from the downstream side to dilute the feed into the last-stage hydrocyclone. This step provides the necessary fluid volume for efficient hydrocyclone operation and helps to remove soluble protein and other solubles in the starch fraction. The overflow, containing protein, solubles, and a small amount of starch, is used to dilute the feed on the previous stage. Recycling of the overflow to the previous stage washes the solubles and protein upstream to the overflow of the first-stage hydrocyclone. The first-stage overflow is recycled to the primary disk-nozzle centrifuges.

The number of hydrocyclone stages in a starch washing system depends upon the ease of separation of starch from protein and the amount of wash water used. Some starches, such as potato starch, require only six to eight stages for adequate washing; other starches, such as corn or wheat, may require up to 14 stages. The number of stages can be reduced if the volume of wash water is increased. About 1.1–1.8 liters of wash water per kilogram of corn (0.13–0.22 gal per pound) is used for maize starch washing.

Because the capacity of individual hydrocyclone (often referred to as a cyclonette) is only 3.7 to 5.7 liters (0.98 to 1.5 gal) per minute, cyclonettes are arranged in parallel in a housing. They share common inlet and outlet manifolds. The most popular assembly, known as a clamshell (**Fig. 11**), contains approximately 480 cyclonettes. Other assembly arrangements are available.

Cyclonettes used in the starch industries are either 0.4 in. (10 mm) in diameter, used for small starch granules such as maize, or 0.6 in. (15 mm) in diameter, used for large starch granules such as potato. Two types of cyclonettes are used in a starch washing system. The A-type cyclonettes are 0.4 in. (10 mm) in diameter with a 0.1-in. (2.5-mm) feed inlet, a 0.1-in.

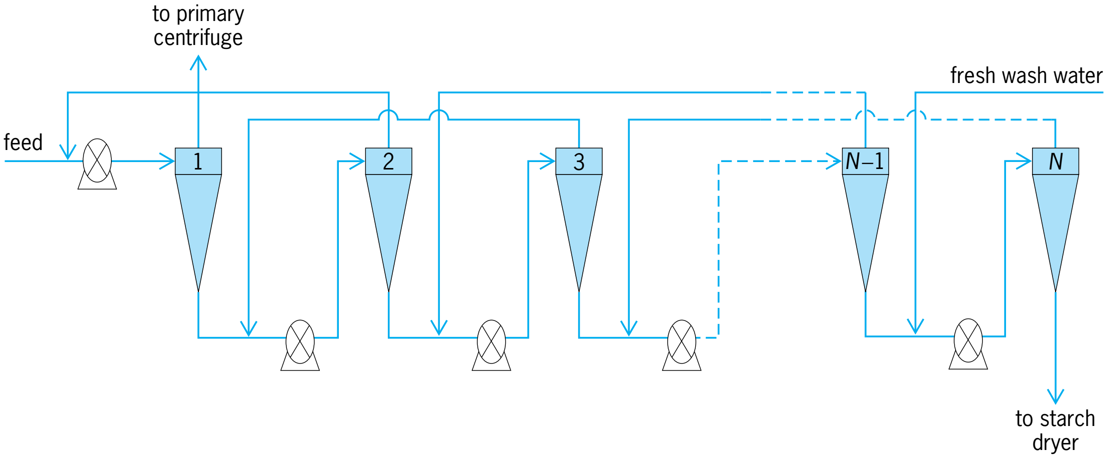


Fig. 10. Starch washing system using a number of hydroclone stages.

(2.5-mm) vortex finder opening, a 0.090-in. (2.3-mm) heavy-fraction opening, and a cone taper of 6°. This is the standard cyclonette used throughout most of the stages of starch washing. Approximately 60% of liquid flow goes to the overflow (lighter fraction) and 40% to the underflow (heavier fraction). The B-type cyclonettes are also 0.4 in. (10 mm) in diameter with similar dimensions, except that the cone taper is 8° and the heavy-fraction opening (apex) is 0.094 in. (2.4 mm) in diameter. About 70% of liquid flow goes to the light fraction, and only 30% goes to the underflow; therefore a higher starch concentration is obtained in the underflow. B cyclonettes are used in the final or last two stages of starch washing. The 15-mm (0.6-in.) cyclonettes are configured in the A or B types as described above.

Hydrocyclones/liquid cyclones have no moving parts. They rely upon slurry velocity to create adequate centrifugal forces to accelerate settling of the heavier solids. Slurry feed pressures in the range 90–140 lb/in.^{2} (620–965 kilopascals) are commonly used in the starch industry for separating starch and protein. Centrifugal separation of starch from protein in hydrocyclones depends not only on density differences but also on other starch characteristics, such as shape and size of granules.

Starch size, shape, and surface characteristics can enhance or retard separation from the smaller protein particles. Larger granules are more easily separated. Because of differences in shape, surface texture, and density of starch granules in various starches, different lower limits of starch granule size may be effectively separated from the protein among different starches. For maize, granules less than 6 micrometers in diameter will have equal probability of going out in the overflow or the underflow. This lower limit is about 1 and 5 μm for rice and wheat, respectively. The probability of a particle going out

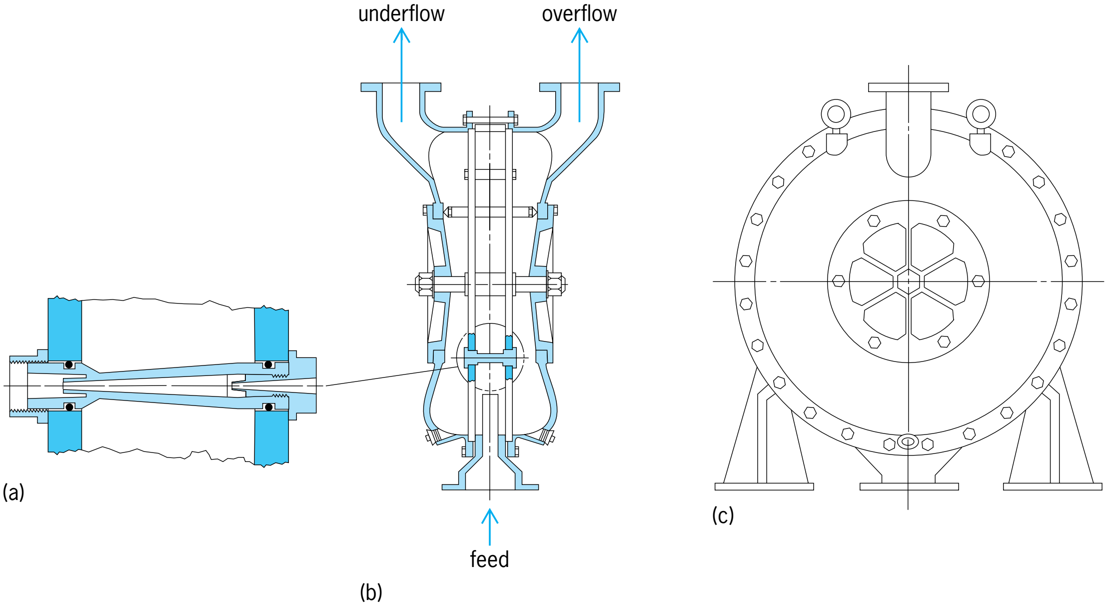


Fig. 11. Clamshell assembly of starch washing cyclonettes. (a) Detail showing cyclonette arrangement. (b) End view showing location of feed, overflow, and underflow streams. (c) Side view of clamshell.

with the underflow decreases with decreasing particle size for all products.

Genetic variation in the starchy raw material, environmental conditions in plant culture, and postharvest drying of the starchy raw material can also affect starch protein separation. Research on 10 United States maize hybrids representing a range of cultivars showed that although the initial and final starch granule size distributions were very similar among all hybrids, the recovery of starch varied widely. There apparently are starch density differences within a hybrid and between hybrids. Uniformity of starch granule size also influences separation. A more homogeneous starch granule distribution results in better separation.

A starch slurry containing approximately 35% starch by weight is commonly used, although the cyclone will efficiently separate more dilute concentrations. Higher concentrations of feed solids are preferred to increase production in the same size of equipment. However, very high concentration may plug the cyclonettes.

Larger hydrocyclones are used to separate germ from endosperm and fiber. Germ separation follows the first two grinding steps in the maize wet-milling process. These hydrocyclones (often referred to as germ clones) release the hydrated germ from the pericarp and endosperm. They are operated at differential pressures of 30–45 lb/in.^{2} (207–310 kPa). The germ clones are commonly 6 in. (152 mm) in diameter and 3.3–3.9 ft (1–1.2 m) long. Larger germ clones [8 in. (203 mm) or 9 in. (229 mm) in diameter] are increasing in popularity because of higher capacity, and they are less prone to plugging because of their larger overflow outlet.

Hydrocyclones may also be used in other solid-liquid separations such as recovery of suspended solids from liquid processing-plant waste effluents and separation of starch, soluble sugars, and fiber from proteins in ground oilseed slurries for protein concentration. In all hydrocyclone applications, the suspended solids must be carefully characterized so that the appropriate size of hydrocyclone and fluid flow rates for effective separation can be determined. Steven Eckhoff

Heating and Cooling

Heat transfer is one of the most common and important unit operations in food processing. It is used to change the temperature of the food product and to change the phases present in the food (vapor, liquid, solid). Thus, preservation processes such as pasteurization, blanching, and canning are based on principles governing heat transfer, just as are drying and freezing processes. Heat-transfer operations are used to either remove or add heat in order to alter the physical, chemical, or storage characteristics of food products. An example of using heat input to change a food product is cooking to alter texture and develop color and flavor. Heat removal is used to change physical form as in the production of ice cream or candies.

Heat transfer by conduction, convection, and radiation are all employed in food processing. In the application of heat transfer, however, it is necessary to know the effects of temperature on rates of change of the physical and chemical characteristics of the food. Knowledge of food science and engineering leads to the optimum design of heat-transfer operations. *See* HEAT TRANSFER.

The food industry relies extensively on heating and cooling processes (either indirect or direct systems) to control the quality and characteristics of food products. These systems are used with a wide variety of heating and cooling media to effect temperature and phase change. In the application of these methods, the safety of the consumer is of paramount importance. Secondary criteria as to the choice of methods include production rate, cost, energy efficiency, and waste generation. Heating and cooling methods applied to foods can be classified as either indirect or direct.

Indirect methods. When these methods are used, the food product does not come in contact with the material absorbing (cooling) or supplying (heating) the heat energy. The heat is transferred to (from) the product through a physical barrier such as a stainless steel plate or a tube wall for heating (cooling). For heating processes, the heating medium is generally hot water or steam, although other heat-transfer liquids or gas flames may be used in high-temperature applications. For cooling processes, the cooling fluid is usually a commercial refrigerant cycled through a mechanical refrigeration unit.

The simplest type of indirect heat exchanger is a kettle which holds the food product. The outside of the kettle may be heated by direct flame (radiation and convection heat transfer), or the kettle may have an outer wall creating a space through which the heating medium can flow. Kettles are used in batch heating processes, and the product is usually mechanically agitated to increase heat transfer and avoid local excessive hot spots and burn-on (fouling).

For larger industrial operations, liquid food products are heated or cooled in continuous-flow heat exchangers. These take a variety of forms, including one with a tube inside another (tubular exchangers) and flow channels created by several parallel, corrugated plates (plate-and-frame heat exchanger). Milk is an example of a liquid food that is heat-treated in either type of heat exchanger prior to packaging and distribution. For more viscous food liquids, such as starch-based puddings and candies and vegetable purees (baby food), a tubular-type exchanger equipped with a rotating axial scraper blade is used. The action of the scraper blade increases heat transfer by removing the layer of liquid immediately adjacent to the wall and thus minimizes burn-on.

Direct methods. Direct heating systems are not as commonly used as indirect systems in the food industry. In cooling applications, direct systems include cryogenic freezing (using, for example, liquid nitrogen, carbon dioxide, or a commercial refrigerant), liquid immersion freezing (that is, using aqueous solutions of glycerol or sodium chloride), and cooling with ice. The systems differ largely in their cost, rate of freezing or cooling, and effect on productbreak quality.

For heating, water, steam, and air are most frequently used as direct heating agents. In each case, the heating fluid must be of food-grade quality. For water, this presents no major problem since potable water must be used in contact with food material. For air, this frequently requires some filtering treatment prior to contact with the food. For steam, this requires that the steam be produced with FDA-approved boiler compounds. Steam which can be used in direct contact with food is called culinary steam and is frequently made by using boiler steam to make steam out of deionized potable water. When steam is used to heat food directly, the condensate that is produced dilutes the food product and it may be necessary to remove this water. Frequently this is accomplished by boiling off the water (flashing) in the cooling operation (for example, exposing the hot product to a partial vacuum). In steam heating, the steam can either be added to the product (a process called steam injection) or the product can be added to a steam atmosphere, a process called steam infusion. Steam injection is the more popular industrial method.

Heat transfer by radiation is best categorized as a direct method. Electromagnetic radiation in the infrared region (1–400 micrometers) is readily absorbed by water and many of the organic constituents of food, resulting in heat transfer. Infrared heating is used extensively in the baking and roasting industry. *See* HEAT RADIATION; INFRARED RADIATION.

Microwave and dielectric heating are accomplished by electromagnetic radiation in the radio-frequency range. In dielectric heating, the product is placed between parallel electrodes, and the charge on the plates is oscillated at 1000–3000 MHz/s. For microwave heating, electromagnetic radiation in the above frequency range is generated in a magnetron tube and transferred to the product located in a cavity where the energy is absorbed. Microwave and dielectric heating have somewhat limited applications because they are expensive. *See* DIELECTRIC HEATING; MICROWAVE.

Daryl B. Lund

Thermal Processes

The heat treatment of food is one of the important processes for the conditioning of food for preservation. This treatment accomplishes many objectives, including inactivation of microorganisms, enzymes, or poisonous compounds and production of desirable chemical or physical changes in foods.

When compared to typical engineering material such as metals or minerals, food is relatively susceptible to thermal degradation. Therefore, to produce nutritionally sound and microbially safe food products, heat treatment should be accurately controlled in order to accomplish its objectives. To assist in this control, the temperature responses of foods subjected to heat treatments are frequently estimated by using empirical or theoretical heat-balance equations.

Heat treatment. This type of processing follows different patterns. It may involve application of heat indirectly, as in a tubular heat exchanger, or it may be accomplished through the direct contact of the heating medium with the food, as in the baking of bread in a hot-air oven. The principal operations involving heat treatment of foods are blanching, preheating, pasteurization, sterilization, cooking, evaporation, and dehydration.

Blanching is a hot-water or steam-scalding treatment of raw foodstuffs to inactive enzymes which might otherwise cause quality deterioration, particularly of flavor, during processing or storage. Most vegetables and some fruits are blanched before canning, freezing, or dehydrating. The commonly used types of equipment are rotary perforated drums, rotary screw conveyors or troughs, and pipe flume blanchers. Water is the predominant heating medium in drum and flume blanchers; steam is used in the screw-type blancher. Temperature is usually 212°F (100°C) or slightly lower. The length of treatment, when preceding freezing, varies from 50 s to 10–11 min.

Preheating is a treatment used immediately before canning to ensure the production of a vacuum in the sealed container and uniform initial food temperature before further thermal treatments (pasteurization or sterilization). When the product cools, the head-space vapor condenses. Foodstuffs of liquid or slurry type are usually preheated in tubular heat exchangers which are equipped with screw conveyors if necessary. Foods of particulate type, submerged in a brine or syrup in the open-top container, are passed through a heated chamber, known as an exhaust box, in which steam is the usual heating medium.

Pasteurization is relatively mild heat treatment of food and involves the application of sufficient thermal energy to inactivate the vegetative cells of microorganisms, molds, yeasts, or enzymes that are harmful to human beings or to food quality, or both. The inactivation of bacterial spores, which are extremely resistant to heat, is not required, since pasteurized food is usually stored at refrigerator temperatures or the chemical composition of the food prohibits the germination of bacterial spores. Typical heating times and temperatures for pasteurization are 30 min at 149°F (65°C) or 15 s at 160°F (72°C). Heat sterilization requires more thermal energy than pasteurization does and is usually accomplished by heating food which is packed and sealed in containers made of metal, glass, or plastic film laminated with aluminum foil. One heat sterilization or pasteurization treatment is divided into heating and cooling phases. During the heating phase, the food is heated for the proper time by applying the heating medium, usually saturated steam or hot water maintained at the appropriate temperature. The cooling phase begins immediately following the heating phase. Since the temperature is at its maximum, or almost maximum, level at the end of the heating phase, it should be lowered as quickly as possible to avoid any unnecessary thermal destruction of the nutritional or sensory quality of the food. *See* PASTEURIZATION.

There are batchwise and continuous heat-sterilization processes which are in commercial use. Widely used for batch processing, a still retort permits food packaging in metal cans, glass jars, or polymeric film pouches with thermal processing without agitation. The heating media are normally steam for food in cans, hot water for food in glasses, and high-temperature water with air overriding pressure or steam-air mixture (typically 75% steam and 25% air) for food in pouches. Pressure above the saturation water pressure is required for pouch processing to counteract pressure buildup inside pouches, since pouch seals do not have the mechanical strength to withstand the pressure difference. Continuous processes are more economical in terms of heat energy, labor, and time than batchwise processes are, but the cost of equipment for continuous processing is generally greater. The continuous food sterilizing system, which is one of the most widely used in the United States, requires only 50% as much steam and 15–40% as much labor per unit of food processed as conventional batchwise systems need (**Fig. 12**).

The hydrostatic system (**Fig. 13**), which originated in Europe, uses about 25% less steam and water per unit of food processed than the conventional batchwise system of processing does.

The temperatures to which food is heated in conventional sterilization processes are dependent on the acidity of the food. A normal temperature range for the heat sterilization of low-acid food (food with pH 4.6 or above and with water activity equal to or above 0.80) is 221–248°F (105–120°C); and a range for high-acid food (for example, fruit) is 180–212°F (82–100°C).

Some heat sterilization methods are called high-temperature short-time (HTST) processes. An HTST process usually consists of two separate heat treatments applied in sequence. One example of these processes is a free-falling film system. Liquid food, usually milk, is preheated first to about 150°F (66°C) and carried to a falling film sterilizer by a product feed line. In the sterilizer, the food is heated as it free-falls in the form of a thin film in 520–540°F (270–280°C) culinary steam environment. The food temperature increases from about 280°F (140°C) to virtually the temperature of the steam in less than 0.3 s. In another method, liquid food with low viscosity is heated by a tubular or plate-and-frame heat exchanger. Scraped-surface heat exchangers (**Fig. 14**) are used to process highly viscous liquid foods. The heated food then passes through a holding tube for at least 3 s before cooling in a vacuum flash chamber down to 149–158°F (65–70°C) as the excess moisture is flashed off. The purpose of the first treatment is to inactivate enzymes, with most HTST processes, by heating food for 5–10 min at 149–185°F (65–85°C), the second to inactivate microorganisms, generally, by heating food for 3–30 s at 257–302°F (125–150°C). There are sound reasons for the requirement of two heat treatments in HTST processes. The rate of thermal inactivation of a biochemical or biological factor increases as the heating temperature increases. High temperatures have a great influence on microbial inactivation; but, in contrast, enzymes are less sensitive to thermal inactivation at temperatures above 248°F (120°C). Therefore, the first low-temperature and second high-temperature heatings are required for enzymic and microbial inactivations, respectively. In HTST processes, food is usually heat-sterilized before being aseptically filled into containers. Therefore, these processes are normally known as aseptic processes. The commercial application of these processes is limited mostly to liquid food since it is difficult to heat particulate foods quickly to the desired high temperatures. Another reason for such a limitation is difficulty in predicting the flow behaviors and temperatures of foods going through an aseptic processing system. However, many academic and industrial researchers are investigating ways to overcome these difficulties. Sample heat exchangers used by the researchers include scraped-surface heat exchangers (Fig. 14) and microwave heating units. Electrical, ohmic resistance heat is a basis for a heat exchanger developed in England. With this

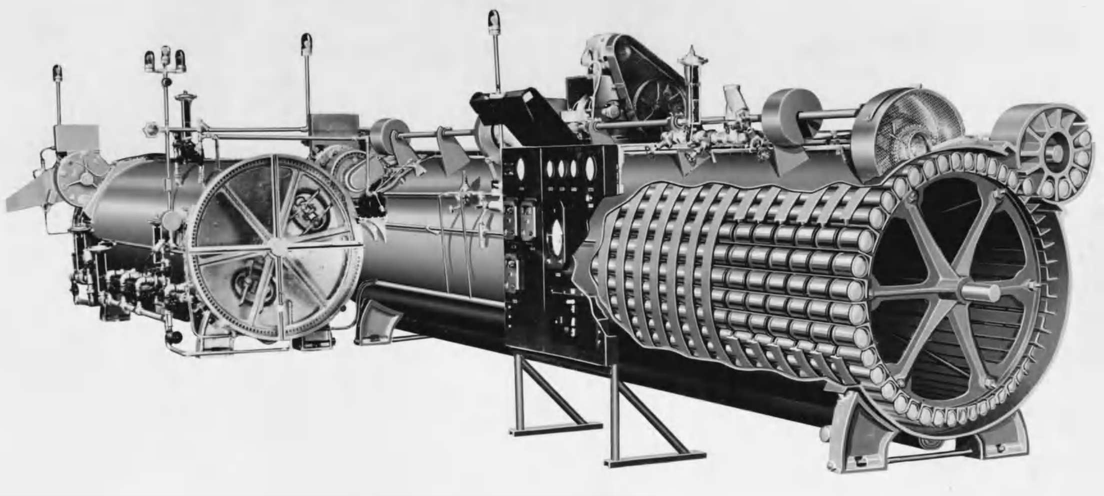


Fig. 12. Cutaway view showing a continuous-pressure food sterilizing system which is of the reel and spiral type. The smaller tank at the left is a continuous pressure cooler. (*FMC Corp.*)

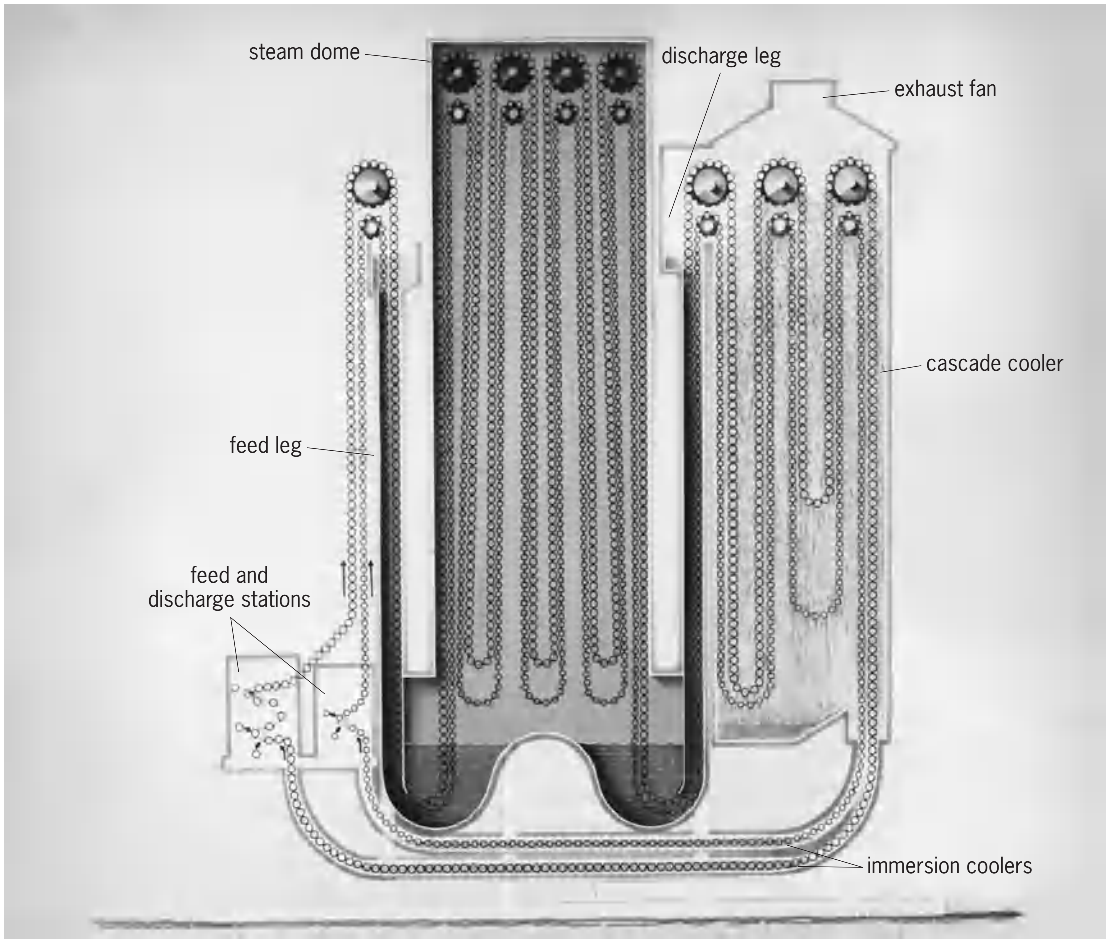


Fig. 13. Hydrostatic system of sterilizing food in sealed containers. (*FMC Corp.*)

device, an alternating current of a proper voltage is fed continuously through food passing within a plastic-lined stainless steel pipe. The pairs of electrodes are placed directly in the foodstream to provide the electric current. The claimed advantages of the ohmic heat exchanger include the ability to heat particulates quickly without disintegration.

Packaged or bulk food undergoes cooling after heating for a proper time at a proper temperature, since maintaining food at a high temperature unnecessarily causes quality deterioration. With a batch process of food packaging, the heating medium in a retort is replaced by cold water for cooling, with air overriding pressure for pouches and glass containers, or without such for most cans. In continuous-process food packaging, the cooling is accomplished by passing the packages in a cold water cooling section (Figs. 12 and 13). In a HTST process of bulk food, a water- or refrigerant-cooled heat exchanger is used. For example, a scraped surface heat exchanger is used for cooling viscous liquid exiting from a heating section.

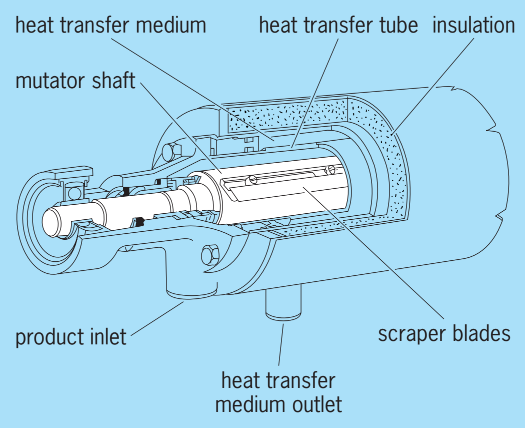


Fig. 14. Votator scraped-surface heat exchanger. (*Chemetron Corp.*)

Batch processes are flexible in operation and provide easy changeover to different products, although the processes are labor intensive. Therefore, batch processes are controlled automatically by computer-based process controllers in several thermal processing plants. Most computer controls are designed to automate starting and ending a process at a controlled temperature according to a standard sterilization process design.

A proper heating time and temperature for food sterilization may be estimated through mathematical combination of food temperature response at the slowest heating point with the thermal inactivation

TABLE 1. Commercial heat treatments for physical or chemical modification of food

Process	Heating medium	Sample food	Heating-medium temperature, °F (°C)
Baking	Direct heating with air and radiative heating with oven walls	Breads, biscuits	240–300 (120–150)
Deep fat frying	Direct heating with oil	Potato chips, noodles, doughnuts	320–365 (160–185)
Drum drying	Indirect heating with steam	Bananas, potato whey	275–293 (135–145)
Evaporating	Indirect heating with steam	Tomato paste	
Extrusion cooking	Direct heating by mechanical energy dissipation through shearing, frequently combined with direct or indirect heating by steam	Breakfast cereals, snack foods, bread substitutes[*]	212–230 (100–110)
Forced-air-drying	Direct heating with air	Vegetables	140–203 (60–95)
Freeze-drying	Indirect heating with heat-exchange fluid or direct heating with thermal radiation	Fish, meat	104–176 (40–80)
Microwave heating	Electromagnetic waves, sample frequencies 915 or 2450 MHz	A variety of foods requiring heating or cooking[†]	
Spray drying	Direct heating with air	Instant coffee, milk powders	401–590 (205–310)

[*] Heat is generated within food through shear (the conversion of mechanical energy to heat). Electrical heaters placed within the barrel of an extruder are used to control changes in food temperature. Typically food temperature at the inlet of an extruder is 212°F (100°C) and increases to 300–360°F (150–180°C) at the exit.

[†] Heat is generated within food through oscillation of bipolar molecules (for example, water molecules) at a very high rate under the influence of microwaves.

kinetic equation of target microorganisms, for example, *Clostridium botulinum* spores of low-acid food processing. Several researchers have developed estimation methods and computer programs for the estimation.

There are available commercial heat treatments whose major objectives are the physical or chemical modification of food rather than the inactivation of microorganisms or enzymes (**Table 1**).

Industrial cooling methods. Since the rate of spoilage and quality deterioration of fresh food is reduced exponentially when food temperature is lowered, food is cooled after slaughtering or harvesting. Methods for postslaughter cooling of carcasses depend on the type of animal. Beef or hog carcasses are cooled in a chilling room with an ammonia-based refrigeration system. Beef carcasses require longer cooling time because of massive size. The chilling room temperatures for beef and hog carcasses are 32–34°F and 37–39°F (0–1°C and 3–4°C) respectively, and the relative humidities are kept high, above 90%, to control weight loss during cooling. The final target temperatures are 35°F (1.7°C) for beef carcasses (average temperature) and 37–39°F (3–4°C) for hog carcasses (internal ham temperature). With beef carcasses, most cooling is done overnight (16–20 h) in a chilling room and completed during subsequent holding-room storage in about 1 day. Hog carcass cooling is completed in 8–12 h in a chilling room. Poultry carcasses are scalded for feather removal and cooled in a continuous slush-ice immersion cooler. With initial carcass temperatures of 75–95°F (24–35°C), the carcasses are cooled to 39°F (4°C) or lower in 1200–2400 s.

Fig. 15. Vacuum cooler of railroad-car size. (*Gay Engineering Corp.*)

Postharvest cooling methods for fresh fruit and vegetables depend on the type and volume of produce and the availability of cooling medium. With hydrocooling, produce is cooled by spraying or immersing in chilled water. For hauling a distance over 100 mi (160 km), hydrocooling is used for freestone peaches, clingstone peaches, tart cherries, cantaloupes, and many vegetables. Forced-air cooling is accomplished by forced flow of chilled air through the produce. For example, grapes are cooled by chilled air of 34.7°F (1.5°C) or lower and of high relative humidity with an airflow rate of 0.0027 ft^3/s lb fruit (0.00017 m^3/s kg fruit). In vacuum cooling (**Fig. 15**), water is forced to evaporate from the surfaces of vegetables or fruits by a vacuum of 29.7 in. Hg (100,292 newtons/m^2) or more created around the product. This system is especially good for products such as lettuce and spinach which have large surfaces in relation to volume. Kan-ichi Hayakawa

Freezing

Freezing is an effective method for preserving many foods. Although some foods lose significant quality when frozen, most food products, if properly frozen and handled, will maintain a fresh condition. The shelf life of frozen foods is extended beyond that of the fresh product because the lower temperature decreases the rate of deterioration, and as the liquid water is changed to a solid, the solutes are immobilized.

Food preservation by freezing is a relatively recent phenomenon. Ice was manufactured nearly 3000 years ago by evaporative cooling and nighttime radiation cooling, but food preservation by freezing did not occur until the 1800s. Prior to the development of mechanical refrigeration in 1851, a limited amount of fish was frozen in salt-ice mixtures. Later, red meats and cheese were frozen for transport. Technical developments in the first half of the twentieth century made rapid growth of the commercial frozen food industry possible. Retail frozen foods have been significant only since 1940.

Water. Liquid water is important to many chemical and microbial deterioration processes in foods. Many foods contain 65–85% water by weight. Frequently some of this water is free to move in the food substance, but contains solutes. This water can be frozen, but the freezing point—the temperature at which the liquid water is in equilibrium with solid water or ice—is decreased by the solutes. Ice crystals contain almost no solute. As a result, when a solution is frozen by decreasing the temperature, pure water is frozen out of the system, increasing the concentration of solutes and decreasing the freezing point of the remaining solution. Thus, free water is frozen over a range of temperatures with the initial freezing point somewhat below the freezing point of pure water.

Many food systems include water that is physically bound to the solids in the food. This bound water usually is not freezable within the normal range of freezing temperatures. However, if the bonds are sufficiently weak, some of these molecules may be frozen in ice crystals (**Fig. 16**).

Ice crystals. As liquid water turns into a solid, it forms into an organized systematic pattern or crystal. Each crystal originates as a number of molecules associating into a tiny ordered particle called a nucleus. After formation, the crystal can grow as additional molecules become associated with the original nucleus. The rates for forming nuclei and for crystal growth both increase with decreasing temperature, but the rate of nucleation is much greater than the rate increase for growth as temperature is decreased. Therefore, very rapid freezing favors the formation of numerous small crystals. During storage the crystal size and shape both change due to recrystallization, even though the product remains below the freezing point.

The location of ice crystals in tissues and cellular suspensions also depends on the rate of freezing. When the temperature decreases slowly (less than 1.8°F/min or 1°C/min), most ice crystals are formed in the locations outside the cells. Liquid water inside the cells diffuses through the cell walls to be frozen in large extracellular crystals. When the freezing occurs rapidly, both intracellular and extracellular ice crystals are formed. For most food products the formation of small ice crystals uniformly throughout the food mass results in the greatest retention of freshlike appearance and food quality.

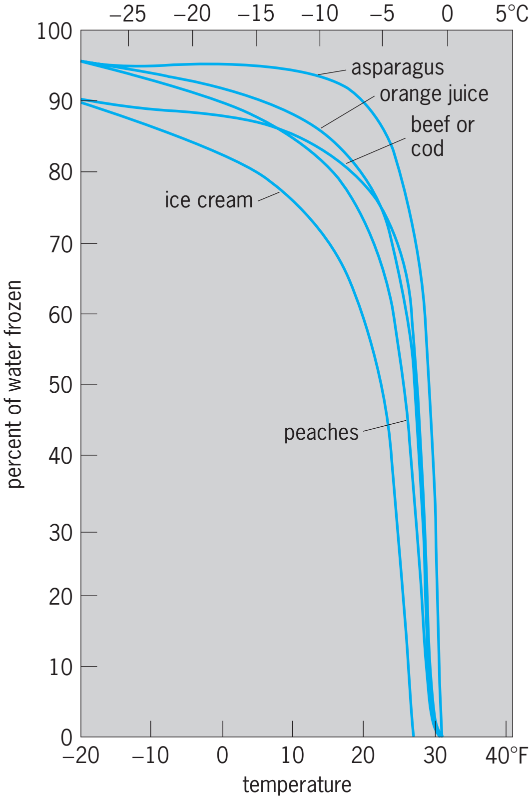


Fig. 16. Percentage of total water that is frozen at any given temperature in several food products.

Quality deterioration. Freezing and frozen storage are effective food preservation methods for most foods, but quality deterioration occurs during freezing, frozen storage, and subsequent thawing. The deterioration rate depends on the enzymes present, the freezing rate, the storage temperature, the consistency of the storage temperature, and the thawing method. An increase in high-quality storage life results when products are stored at lower steady temperatures (**Fig. 17**). The actual increases achieved depend on good packaging that excludes air and reduces the moisture loss from the product.

Most fresh fruits and vegetables contain enzymes that catalyze deterioration reactions during frozen storage. Many deterioration reactions can occur, but the oxidation of various food components is frequently the most important cause of quality deterioration. The naturally occurring enzymes that catalyze these reactions are generally inactivated by blanching (a heat treatment) the food before it is frozen. *See* ENZYME.

Fluctuating storage temperatures generally cause quality losses at rates greater than would be expected for the mean storage temperature. Both recrystallization and quality losses associated with chemical reactions are accelerated by varying temperatures. Small

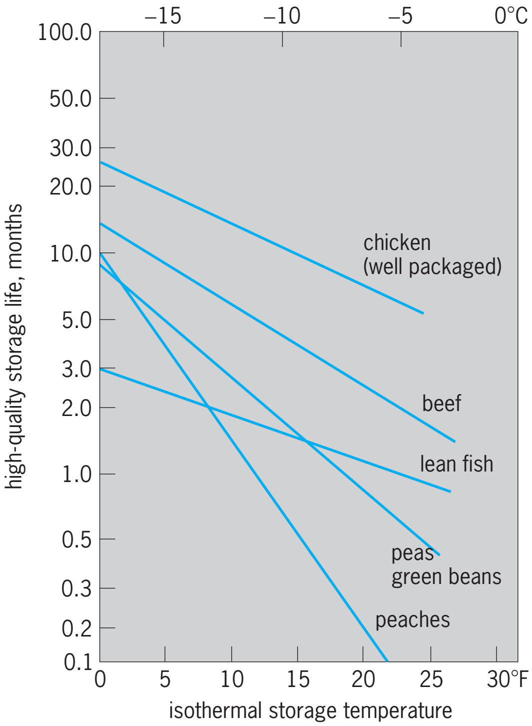


Fig. 17. Influence of isothermal storage temperatures on high-quality product life.

variations inevitably occur in any storage, but the most serious problems occur when the product is being moved from the freezer to the warehouse, between warehouses, or from the warehouses to the point of final consumption.

Thawing is inherently slower than freezing unless the thawing process is assisted by microwave energy or combined with the cooking process. Frequently the temperature gradient (ambient temperature in excess of the freezing temperature) is less than the gradient for freezing. More importantly, the rate of heat transfer through the thawed material on the outside of a food product is much slower than that through the outer frozen shell during freezing. (Of course, these comments do not apply if the product is a fluid that is allowed to flow away from the surface of the frozen core as it thaws.)

For many products there is evidence of accelerated quality loss rates when 10–75% of their water is frozen. In this range, much of the liquid water has been removed as ice crystals. Thus the remaining liquid contains more concentrated reactants that cause accelerated quality losses.

Freezing systems. Numerous commercial freezing methods are available. Each method has its own characteristics that make it more desirable for some applications and less desirable for others.

Freezing in refrigerated air. Many freezing systems use a mechanical refrigerating system to supply cold refrigerant to a set of coils that cool air that has passed through or over the food products. Thus the heat is removed from the product by cold air, and the heat is removed from the air by a mechanical refrigeration system. Freezers in homes and some commercial freezers circulate the cold air through a room or cabinet in which the food is temporarily stored. Typical air temperatures in these freezers are 0 to −22°F (−18 to −30°C). Commonly referred to as sharp freezing, it generally takes several hours to several days depending on the size of the individual food packages. In most freezers of this type the food is stationary during the freezing process, but in a few very large installations the food may move through the freezer on an intermittent or slow-speed conveyor.

Most commercial freezers using refrigerated air force the air over or through the product at a high velocity. Typical air velocities are between 100 and 3500 ft/min (30 and 1000 m/min). These freezers are generally referred to as air-blast units. In most installations the food is moved through these freezers by conveyors. The conveyors may be in a straight line, stacked above one another, or arranged as an inclined spiral. Either packaged foods or individual food particles can be frozen in these freezers. Particulate foods such as peas, beans, or shrimp are frequently frozen in a freezer in which the air is forced up through the conveyor belt and the product. The velocity is sufficiently high to create a fluidized bed, lift the product from the conveyor belt, and keep it agitated, so that each particle is individually frozen. These freezers are referred to as individually quick-frozen units. Air temperatures in air-blast freezers are generally maintained between −22 and −40°F (−30 and −40°C).

In another method of refrigerated air freezing, the air is cooled by compression, heat exchange, and expansion. The air that will remove the heat from the food is also the refrigerant. This system is designed to cool the air to −250°F (−150°C), which means it is also considered to be a cryogenic freezer. *See* CRYOGENICS.

Immersion freezers. Because the heat transfer rate between air and a solid is relatively low, many freezers are designed to contact cold liquids or solids with the food products. The product may be immersed in the cold solution, mixed with a cold powder, or sprayed with a cold solution. Salt brines and sugar solutions have been used since the early twentieth century to quick-freeze fish products and, more recently, canned citrus juice. Many novelty products are frozen in molds that are immersed in a refrigerated solution. Some products (notably, sweet corn on the cob) were frozen by a combination of spray and direct immersion in commercial chlorofluorocarbon refrigerants; these freezers are being phased out because of the environmental hazard of these refrigerants.

Both liquid nitrogen and liquid carbon dioxide are sprayed directly on food products for very rapid freezing without the capital expense of a large mechanical refrigeration system. At atmospheric pressure, liquid nitrogen boils at −320°F (−196°C). Both the vaporization of the liquid nitrogen and the sensible heat required to raise the temperature of that vapor to approximately the food temperature is used to remove heat from the food product. Carbon dioxide exists as a liquid only under pressure. When the liquid is sprayed on a food product, it changes state to dry ice or snow at −109°F (−79°C). The snow absorbs heat and sublimes to a vapor. Both the heat of sublimation and the sensible heat to warm the vapor to near the product temperature are available for freezing the food. Freezing with either of these liquids is considered cryogenic because the phase change occurs below −76°F (−60°C).

Freezing with direct immersion or spray is generally very quick. Even for relatively large food products or packages the time is generally under an hour, and for smaller products it may require only a few minutes.

Indirect freezant-contact freezers. When freezing packaged products, it is important to maintain the intended package geometry. For example, rectangular packages tend to swell, thereby making them difficult to handle or case. By placing the unfrozen packages between refrigerated metal plates that are spaced to the intended package thickness, the dimensional characteristics can be maintained during freezing. The heat from the food is extracted through the metal surfaces and into the brine or refrigerant circulating inside the plates. Nonpackaged foods are also sometimes frozen in indirect-contact freezers using either flat plates or shaped cells. The most common indirect-contact freezers are plate freezers (with the plates stacked either vertically or horizontally and with either automatic or manual product handling), ice cream freezers (either continuous or batch style), and the novelty-product freezers (for both stick and stickless novelty items).

Heat transfer by conduction is relatively rapid, so the freezing times are typically between 15 min and 2 h, depending on the size of the item.

Refrigeration requirements. The energy that must be removed to freeze food is substantial. The refrigeration requirements include this heat plus energy losses from the freezer and refrigerating system, the heat that must be removed from any packaging materials, and heat added by product handling equipment and system inefficiencies. Typically, the food enters the freezer at a temperature above the freezing point (the temperature at which water is first crystallized in the product). Before freezing can begin, the sensible heat (the heat associated with a decrease in temperature) must be removed to bring the product to the freezing point. Typically the product temperature drops below the freezing point and then warms to the freezing point. This dip in the temperature curve (**Fig. 18**) is caused by supercooling (there are insufficient ice crystal nuclei available for growth to provide the latent heat of crystallization at the rate it is being removed by the freezer). At the supercooled temperature, additional nuclei form rapidly, and the rate of crystal growth is limited only by temperature and the rate of heat removal by the freezer. Throughout the freezing zone, the freezer is removing primarily the latent heat for crystallizing the water. There is also a small amount of sensible heat removed because of the temperature decreases associated with greater water removal (Fig. 16). In the postcooling zone, most of the freezable water has been crystallized, and the heat removal is predominantly sensible heat in the frozen product. The solid components in the food do not change phase, and therefore only the sensible heat must be removed from them.

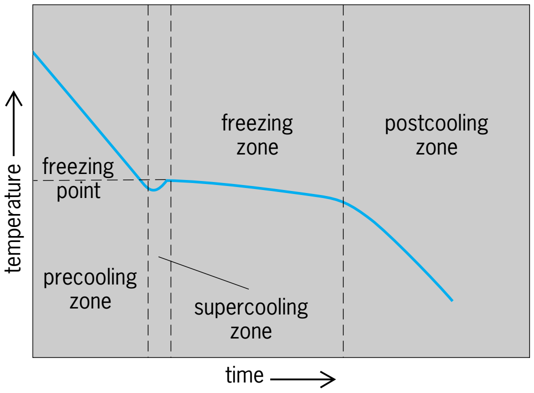


Fig. 18. Temperature profile for food freezing, showing heat transfer regimes. (*After S. D. Holdsworth, Physical and engineering aspects of food freezing, in S. N. Thorne, ed., Developments in Food Preservation, no. 4, 1987*)

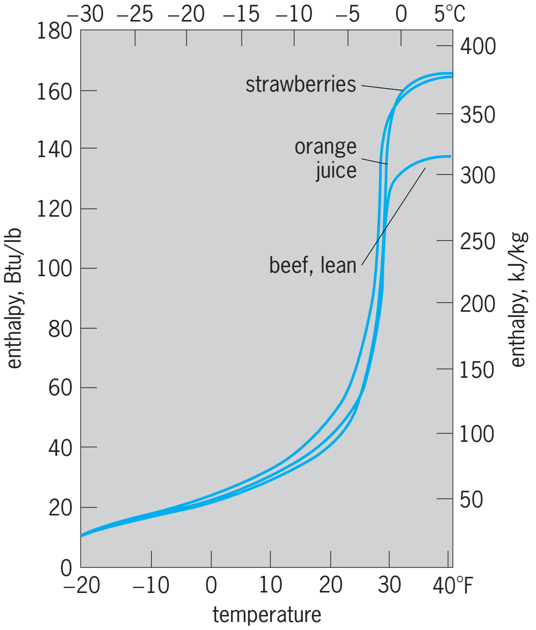


Fig. 19. Energy requirements for food freezing.

With a knowledge of the specific heats of the solid components in the food, of liquid water, and of ice, the latent heat for changing liquid water to ice, and the curve (Fig. 16) for the product being frozen, it is possible to calculate the energy requirements for freezing a product. Alternatively, the enthalpy-temperature curve (**Fig. 19**) for many products has been measured experimentally. The energy that must be removed to freeze a product is equal to the enthalpy difference between the initial and final temperatures of the product. *See* ENTHALPY; REFRIGERATION; SPECIFIC HEAT.

David R. Thompson

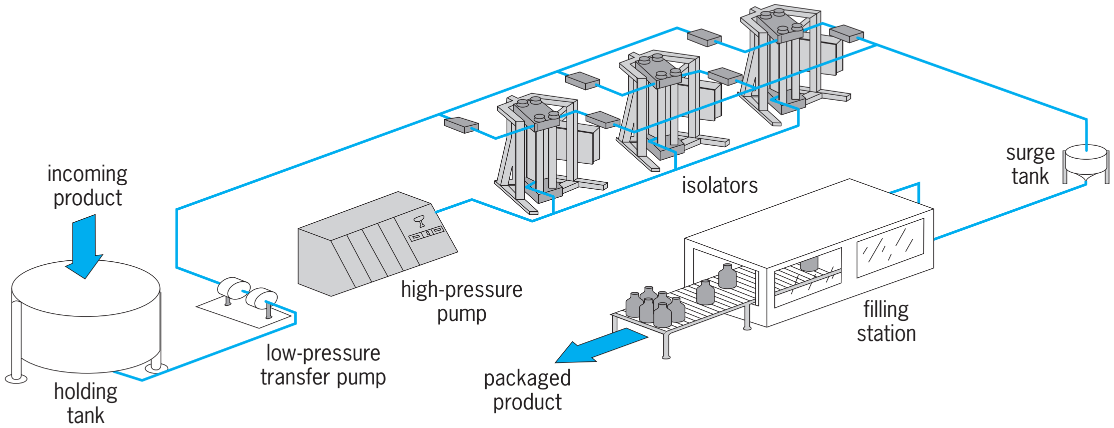


Fig. 20. Schematic diagram for a semicontinuous high-pressure operation. (*ABB Autoclave Systems, Inc.; Flow International*)

High-Pressure Processing

High-pressure (60,000–150,000 lb·in^{-2}; 400–1000 MPa) processing is a nonthermal method used to denature proteins and inactivate microorganisms and some enzymes in foods.

Advantages. High-pressure processing has many advantages over conventional methods. Since it has no effect on covalent bonds, it helps to preserve natural flavors, colors, and nutrients in foods. However, functional properties of food components, such as proteins and polysaccharides, can be altered. Due to its instantaneous and uniform distribution, it can be applied independent of the size and geometry of the product. Consequently, processing times are short, generally in the order of minutes. Once the system is under pressure, there is no loss of mechanical energy. The phase-transition behavior of water can be used to design unique freezing and thawing applications for foods. In general, liquids are used as working fluids; and unlike gases under pressure, liquids are much safer.

Disadvantages. The capital costs of high-pressure processing are considerably more than conventional methods. In addition, spores of some microorganisms are not inactivated, as well as some enzymes. Due to the compressibility of the working liquids, there is generation of heat during compression. Depending on the pressure level, this heat can be significant and void the nonthermal advantage of this process. Gases in the package are undesirable and may cause distortion of the package. At present, most commercial high-pressure operations are limited to the batch type. There is a lack of information

TABLE 2. Commercially available food products treated with ultrahigh pressure

Product	Pressure, MPa	Time, min	Temp., °C	Purpose
Fruit-based: jams (apple, kiwi, strawberries); jellies; purees; yogurts; sauces	400	10–30	20	Pasteurization, improved gelation, faster sugar penetration, residual enzyme activity
Grapefruit juice	200	10–15	5	Reduced bitterness
Mandarin juice	300–400	2–3	20	Reduced odor of dimethyl sulfide, reduced thermal degradation of methyl methionine sulfoxide, replaces first thermal pasteurization
Sugar-impregnated tropical fruits (kept at −18°C without freezing)	50–200			Faster sugar penetration and water removal
Raw ham	250	180	20	Faster maturation, faster tenderization, effect on microorganisms, improved water retention
Raw squid + salt + autolysis, raw scallops				Microbial sanitation, tenderization, control of autolysis
Fish sausages, terrines, and puddings	400			Gelation, microbial sanitation, good texture of raw gel
Raw sake (rice wine)				Yeast inactivation, fermentation stopped
Mochi rice cake, Yomogi fresh aromatic herbs, hypoallergenic precooked rice	400	10	45 or 70	Microbial reduction, enhances rice porosity and salt extraction of allergenic proteins
Ice-nucleating bacteria (used for fruit juice and milk concentration)				Inactivation of *Xanthomonas* bacteria by pressure without loss of ice-nucleating properties
Spiced mashed avocado (guacamole)	400–700			Pasteurization
Fruit juices, yogurt, and acidic foods	400–700			Pasteurization

regarding high-pressure kinetics of enzyme or microbial inactivation, and quality/functionality retention or deterioration. There is also insufficient information regarding the mechanisms involved in the application of high pressure. This includes the requirement to understand, control, and optimize pressure effects on a cellular/molecular level. In addition, more insight is needed to understand the interactions between food constituents under high pressure.

Treatment. In general, treatment with high or ultrahigh pressure (UHP) is accomplished by hydrostatic pressure; that is, the high pressure is indirectly applied to the food, sealed in a flexible or semirigid package, by the use of a working fluid. The process can be batch or semibatch.

The batch process is currently the preferred method for food treatment with high pressure, mainly to minimize problems related to hygiene, flexibility, and technical feasibility. In batch processing, the food is sealed in a hermetic flexible package and put in a pressure chamber, typically a cylindrical pressure vessel made of low-alloy steel, in which pressure is applied through the pressure-transmitting liquid. The production cycle includes the steps of loading the product into the chamber, pressurizing and depressurizing cycles, and unloading of the product. The work of compression can be delivered directly or indirectly to the working liquid. In the case of direct application of pressure, the working liquid is compressed by a piston driven into the cylindrical chamber by a low-pressure pump at one end, taking advantage of the hydraulic principle. This allows a short compression-cycle time, but the operation is limited by the high-pressure dynamic seal between the piston and the internal vessel surface. This type of operation is generally limited to small-capacity systems. The indirect compression method uses a high-pressure intensifier pump to force the working liquid into the pressure vessel until the desired pressure is reached. In this case, only a static high-pressure seal is needed.

By using several pressure vessels and alternating their cycles, a semicontinuous operation can be obtained (**Fig. 20**). In this case, while a vessel is being loaded, another vessel is kept under pressure, and still another is unloaded. Alternatively, pressure can be accomplished by the use of an aseptic continuous system. Current commercial applications of this technology are shown in **Table 2**. *See* HIGH-PRESSURE PROCESSES. Murat O. Balaban; Roger D. Barbosa

High-Intensity Pulsed Electric Fields

Pulsed electric fields (PEFs) is an emerging nonthermal food preservation technology used to pasteurize and sterilize pumpable liquid foods (such as milk, fruit juices, liquid whole egg, and soups) without using lethal heat, ionizing radiation, or chemicals. Products are minimally affected by PEF processing. Very short pulses of high-intensity electrical fields produced by a high-voltage source can significantly reduce the level of pathogens as well as spoilage microorganisms in just a fraction of a second. PEF processing is considered superior to thermal processing because it avoids or greatly reduces the detrimental changes to the sensory, physical, and nutritional properties of food caused by thermal processing. The application of PEF technology to food preservation provides an alternative to thermal processing by producing freshlike products with high nutritional qualities, no additives, and longer shelf lives. Furthermore, PEF assists in reducing preheating and posttreatment cooling of some food products; therefore, energy consumption is considerably reduced compared to conventional thermal processes.

Process. During the PEF process, microorganisms are inactivated by irreversible structural changes in the membrane, leading to formation of channels or pores and destruction of the semipermeable barrier of the membrane. The lethality of PEFs on microorganisms is caused by the action of high-intensity electric fields, and is not due to temperature effects. The high-intensity PEFs cause disruption, permeabilization, and lysis of the cell membrane. This bactericidal effect is related to the electric field intensity, treatment time, and number of pulses, frequency, and pulse width. Microbial inactivation with PEFs is also dependent on the type of microorganisms, microbial growth stage, initial concentration of microorganisms, medium conductivity, and ionic strength. The basic components of a continuous PEF system consist of a high-voltage power supply; an energy storage capacitor; a treatment chamber; a pump to conduct food through the treatment chamber; a cooling device; voltage, current, and temperature measurement devices; and a computer to control operations (**Fig. 21**).

The types of high-voltage power supplies that may be used to charge the capacitor are generated by an ordinary source of direct current (dc). Alternating-current (ac) power from the utility line (60 Hz) is transformed into high-voltage ac power and then rectified to a high-voltage dc. Another way to generate high voltages is with a capacitor-charging power supply which uses high-frequency ac inputs and provides a command charge with higher repetitive rates

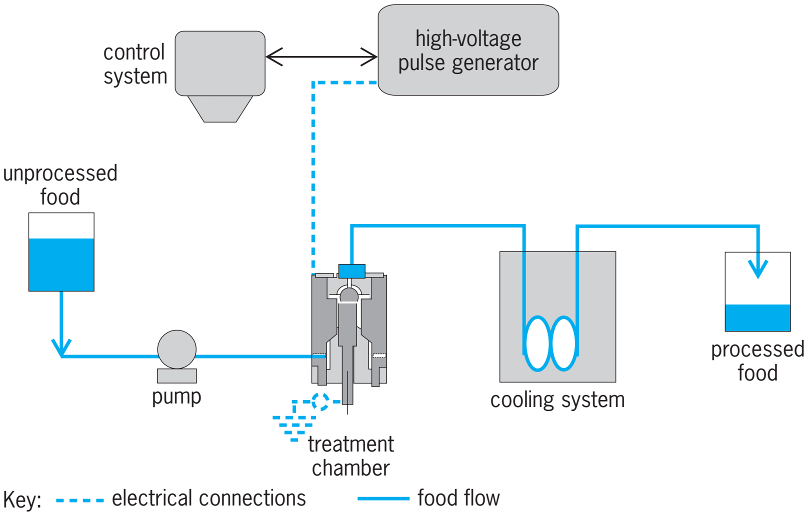


Fig. 21. Continuous pulsed electric field (PEF) processing system.

than a dc power supply. Energy from the high-voltage power supply is stored in the capacitor and discharged through the food material to generate the necessary electrical field in the food.

The energy stored in the capacitor can be discharged almost instantaneously (in a millionth of a second) at very high levels of power. The discharge is accomplished by using high-voltage switches that must be able to operate reliably at a high power and repetition rate. The type of switches can be selected from gas spark gaps, vacuum spark gaps, solid-state switches, thyratrons, and high-vacuum tubes.

The treatment chamber is used to transfer high-intensity PEFs to foods. Cooling water is recirculated through the electrodes, thus controlling the temperature of a sample in the gap formed by the two electrodes in the treatment chamber. A variable speed pump regulates the flow rate of the food during treatment. Chambers with uniform electric fields subject all bacterial cells to the same potential. This is advantageous for electroporation, where no alignment or pearl-chaining is required, and can result in high yields provided the field strength is set at an optimum value. However, since the yield might drop substantially above or below the optimum field strength, the optimum value is not always known. A practical realization of uniform field chambers is a chamber with flat parallel electrodes or a coaxial chamber with a diameter many times larger than the electrode gap. Chambers with nonuniform fields allow alignment at low cell densities with an ac field for fusion and reduce the need to have exactly the right field strength to obtain high transfection yields.

Electrical parameters such as voltage and current waveforms applied to fluid foods can be recorded via a digital data acquisition system. A digital storage oscilloscope records outputs from voltage and current monitors. The computer and oscilloscope should be placed in a shielded area to minimize electromagnetic interference, and electrical and flow parameters should be selected so that each unit volume of the food is subjected to the necessary number of pulses to create the desired inactivation of microbial cells.

After processing, the food is packaged into sterile containers. It is extremely critical to maintain a contaminant-free environment during processing and packaging to avoid cross-contamination of the product. All of the equipment and work area must be thoroughly cleaned and sanitized before processing, and the personnel involved in the operation should be properly equipped to maintain a clean working area.

Application in food preservation. The process of subjecting a food to electric fields may be accompanied by the generation of heat. If electric fields are applied to a food product in the form of short-duration pulses rather than passing electricity directly through it, minimal heat is generated and the process remains nonthermal. Since the pulse repetition rate also contributes to heat generation, the processing system generally includes a means to provide cooling of the treatment chamber, which is one of the most complicated components of the processing systems. Usually, static chambers are used for preliminary laboratory experiments, while continuous chambers are used in pilot plant and industrial operations.

Due to the advantages that PEF presents as a nonthermal process, products such as fruit juices, milk, yogurt, beaten eggs, and tap water have been under extensive research in order to implement the process at an industrial level. Flavor freshness, low energy utilization, and extended shelf life are some of the virtues attributed to PEF treatment. It has also been demonstrated that the textural and functional attributes of some products can be preserved at safe microbiological levels. **Table 3** lists food treated with PEF as well as the process conditions used.

PEF-treated beaten eggs are just one example of the appeal this technology has, since egg products processed by traditional heat treatments exhibit operational problems due to protein coagulation in heat exchangers. However, since PEF enables microbial inactivation at lower temperatures, it represents an important alternative to the egg industry. Significant shelf-life extensions with minimal changes in the physical and chemical properties of certain other foods have also been demonstrated. Experimental results have shown that, in general, the sensory properties of foods are not degraded by PEF. In fact, electric-field-treated products, such as green pea cream soup and fluid eggs, were preferred by panelists in sensory tests over at least one commercial brand product.

Electric-field treatment is also associated with energy efficiency. In the treatment of apple juice, energy utilized with the PEF technology is 90% less than the amount needed for the high-temperature short-time (HTST) processing method. *See* ELECTRIC FIELD.

Gustavo V. Barbosa-Cánovas; Juan J. Fernández-Molina; Barry G. Swanson

TABLE 3. Liquid foods processed with pulsed electric field (PEF)

Food product	Electric field kV/cm	Pulse duration, μs	Pulse number	Treatment temperature, °C
Fresh orange juice	35–60	0.85–1.46	2.5–3.3	25
Cranberry juice	39	4	50	30
Fresh apple juice	50	2	16	45
Skim milk	30–40	2	30	28
Raw milk	40	2	20	50
Green pea soup	35	2	32	53
Liquid whole eggs	40	2	50	30

Concentration

Water-rich liquid foods are concentrated to (1) provide storage stability; (2) facilitate subsequent drying; (3) produce saturated solutions and cause crystallization; (4) produce supersaturated solutions which will form glassy or amorphous solids; (5) reduce storage, shipment, and packaging volumes and costs; and (6) induce flavor and texture changes. Liquid foods are usually concentrated by evaporation; but freeze concentration, reverse osmosis, and ultrafiltration are also used.

Evaporation. Food evaporators usually contain heat-transfer tubes surrounded by a steam- or vapor-filled shell (**Fig. 22**). The liquid to be concentrated flows through the tubes. Part of the water in the liquid is vaporized by the heat that is provided by the steam or vapor (which condenses outside the tubes). The water vapor and remaining liquid separate in a settling chamber or cyclone and leave the evaporator through separate lines.

In most cases, groups of evaporators are interconnected to form multistage, multieffect systems (**Fig. 23**). Liquid feed enters the first stage, and concentrated product is discharged from the last stage. Partially concentrated liquid is pumped out of the first stage into the second stage, and similar transfers occur in the remaining stages. Steam or recompressed vapor is fed into the first-effect steam chest, and the vapor produced flows into the second-effect steam chest. Similar transfers occur in the remaining effects. The vapor from the last effect is condensed externally, and the condensate is discharged through a barometric leg. Steam jet ejectors or a liquid-ring vacuum pump are used to remove noncondensables, which interfere with heat transfer, from evaporator systems.

The evaporation in successive effects of a multieffect system occurs at progressively lower pressures and, consequently, at progressively lower temperatures. This provides temperature differences, which cause heat transfer from the condensing vapor to the evaporating liquid. The heat provided by vapor condensation generates new vapor from the evaporating liquid, which in turn generates more vapor in the following effect, and so on. In a well-designed, multieffect system, the amount of water evaporated roughly equals the amount of steam used in the first effect times the number of effects. However, if steam is used to heat the liquid feed to its boiling point, less than this number of unit weights of water will be evaporated. If the feed is preheated by using vapor from the effects (as shown in Fig. 23), the weight of water evaporated will be very close to the number of effects times the weight of steam used. In beet sugar refineries, vapor is removed from evaporator effects to provide process heating. A smaller ratio of vapor to steam used results, but overall energy consumption in the refinery is reduced.

When evaporation conditions are stable, high-pressure steam passed through an ejector can be used to compress part of the vapor produced in a multieffect system. The compressed vapor and steam provide the heat used in the first effect. When vapor from effect m in an n-effect system is compressed by using this arrangement, $n + m$ unit weights of evaporation are obtained per unit weight of steam used.

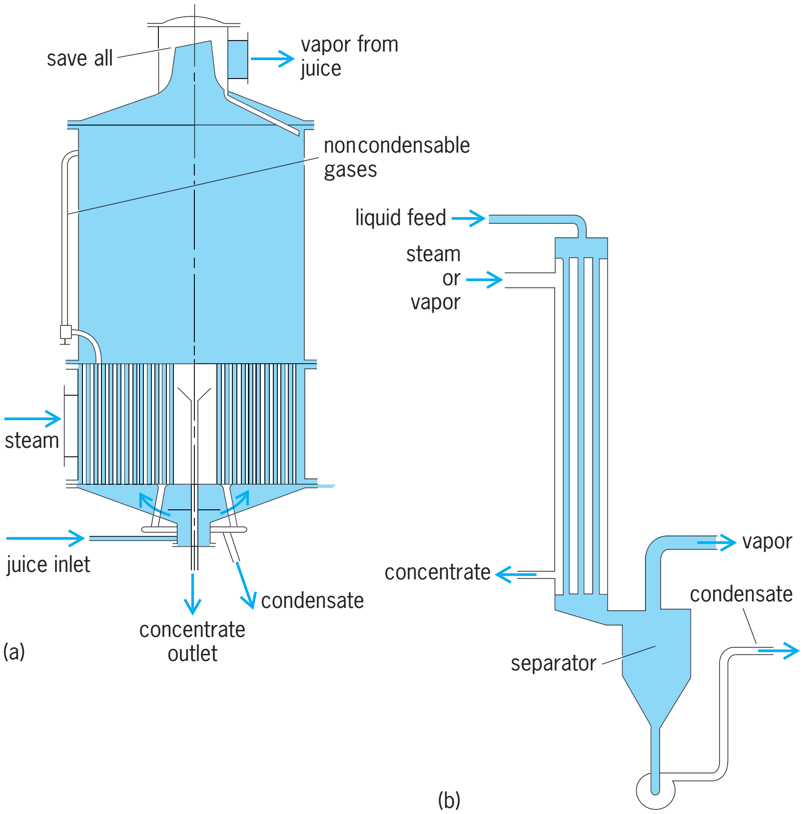


Fig. 22. Single-effect food evaporators: (*a*) short-tube type, (*b*) long-tube type.

Relatively low first-effect temperatures (for example, 162°F or 72°C for skim milk) must be used to prevent product degradation or heat-transfer surface fouling. Temperatures in the last effect, which are controlled by the condensation temperature of the vapor leaving the system, usually range between 104 and 113°F (40 and 45°C). The multieffect system is driven by the difference between the first-effect steam temperature, minus the sum of the boiling point elevations in the effects, and the evaporation temperature in the last effect. The temperature difference per effect is directly proportional to this temperature difference and inversely proportional to the number of effects used. Hence, the total heat-transfer area required in a multieffect system is directly proportional to the number of effects.

Mechanical vapor recompression. This process has been increasingly used to raise the pressures of vapors generated in multistage and multieffect systems. Either the compressed vapors are sent back to the shells of all stages in parallel (**Fig. 24*a***), or vapor from the last effect is sent back to the shell of the first effect (in Fig. 24*b*). Parallel stages are sometimes mounted in a single shell to reduce equipment costs. Very large centrifugal compressors are usually used to handle the vapor flows generated in food evaporators. The maximum temperature increase which

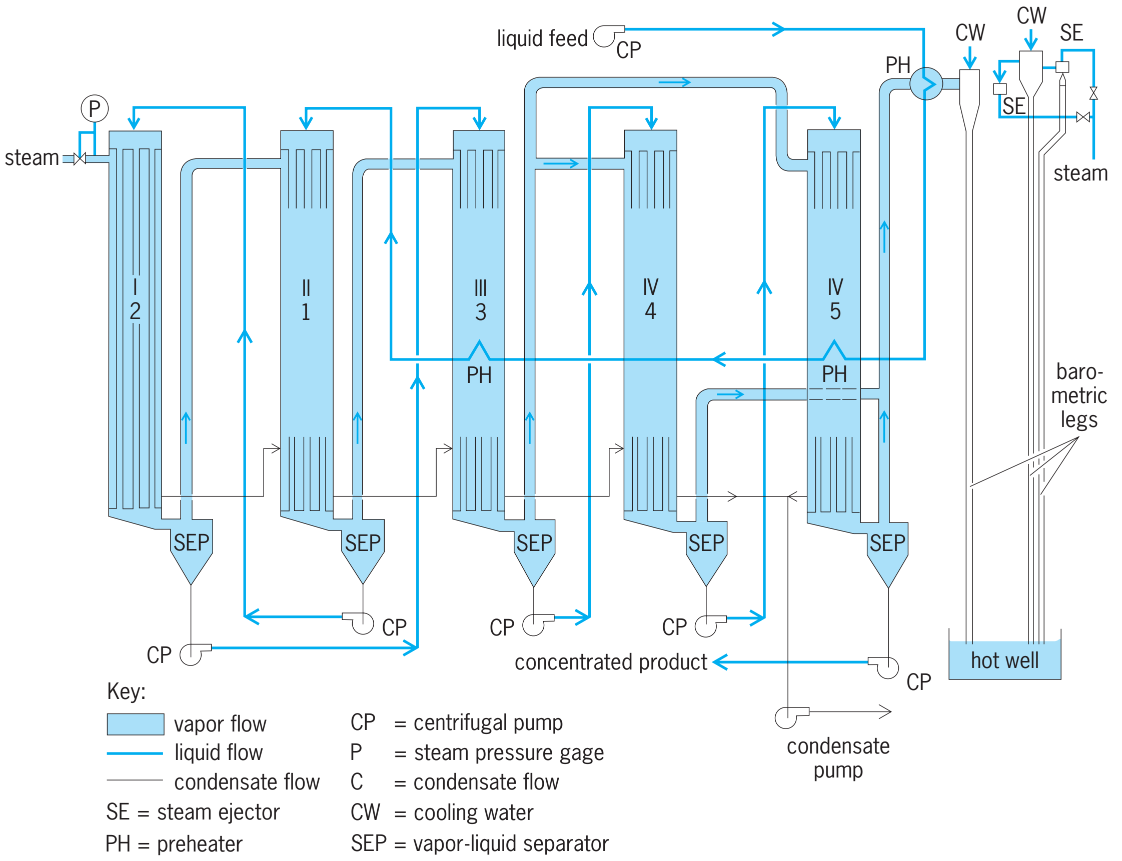


Fig. 23. Multiple-effect evaporation system. Stage numbers (arabic) denote the direction of liquid transfer, and effect numbers (roman) denote the direction of vapor transfer. The first stage is in the second effect, and the fourth and fifth stages are in the fourth effect.

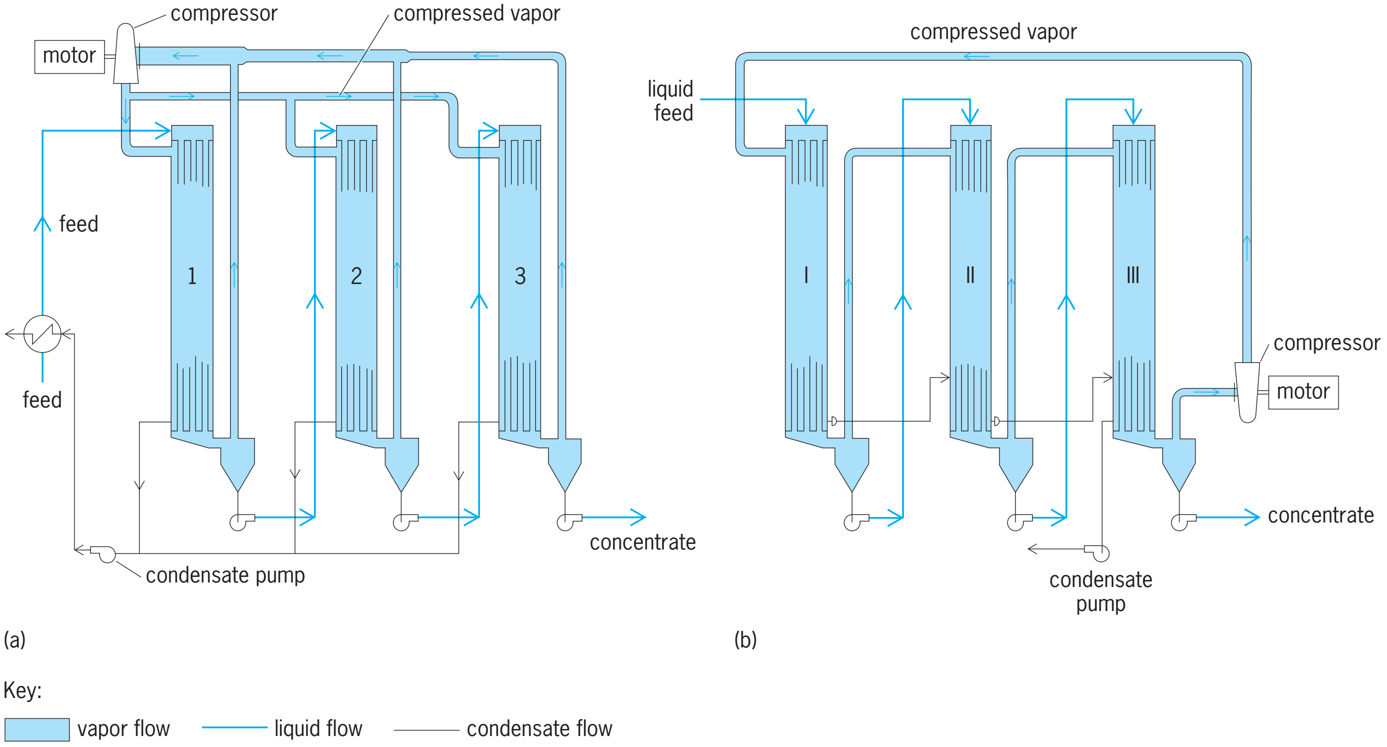


Fig. 24. Vapor-recompression evaporation systems: (*a*) multistage, parallel-mechanical type (arabic numerals are stage numbers); (*b*) multiple-effect mechanical type (roman numerals are effect numbers).

such compressors can provide ranges from 23 to 29°F (13 to 16°C). Lightweight, low-compression-ratio, centrifugal blowers that provide small temperature rises but very high thermal efficiency are used in large parallel-stage systems. In small systems, double-lobe blowers, which can provide larger temperature increases, are used. The energy used to drive a mechanical-vapor-recompression compressor providing a 23°F (13°C) temperature difference roughly equals the heat that would be used in a steam-driven, 17-effect system. However, in terms of the fuel energy required to produce the mechanical energy which drives the compressor, the mechanical-vapor-recompression system corresponds to a 7.5-effect system.

Tubes. Long, vertical tubes (Fig. 22*a*) are used to facilitate single-pass operation in most modern evaporators. Liquid feed, which is pumped into the tops of the evaporator stages, flows in falling films down the inside walls of the tubes. Single-pass, falling-film operation tends to minimize product holdup time and degradation. Falling-film operation also provides the unimpeded vapor flow needed in mechanical-vapor-recompression systems.

Arrays of short vertical tubes, which pass through a steam- or vapor-filled shell called a calandria (Fig. 22*b*), are used in most older evaporators. Many evaporators of this type are used in sugar refineries. Pools of liquid fill the bottoms of such evaporators and barely cover the tubes. Natural convection, promoted by vapor bubble formation in the tubes, causes liquid to recirculate through the tubes and a large-bore downcomer. Because of this circulation, liquid-holdup times are relatively long, and significant product degradation (for example, inversion) occurs.

Instead of vertical tubes, horizontal or tilted tubes and plate- and panel-type heat-exchange surfaces are sometimes used in evaporators. Forced, multipass circulation is used to prevent fouling and facilitate heat transfer when viscous products, such as tomato paste, are concentrated. Scraped heat-transfer surfaces are often used when very heat-sensitive or highly viscous products are concentrated. Rapidly rotating, conical heat-transfer surfaces are also used for such products. Concentrations in the 90% range can be obtained, for example, for candymaking, by using slightly pitched scrapers that both provide thin films and propel concentrates toward the discharge end of the evaporator. The concentrates obtained are so viscous that an expeller is used to pump them out of the evaporator; a large part of the heat used for evaporation is produced by friction.

Aromas and volatile flavors are invariably lost when fruit juices are concentrated by evaporation. Aromas can be recovered from vent vapor leaving the steam chest following the first effect. Carryover of aroma in vent vapor is improved by using countercurrent condensation in that steam chest. Aroma losses can be almost completely avoided by using freeze concentration instead of evaporation. *See* EVAPORATION; EVAPORATOR.

Freeze concentration. In freeze concentration units, liquid feed passes through a refrigerated, scraped-surface heat exchanger, where part of the water contained in the feed is converted into small ice crystals. This concentrates the remaining liquid. Centrifuges are sometimes used to directly separate the concentrate from the ice, but clean separations are not achieved when the slurry sent to the centrifuge contains small crystals. Therefore, the ice-concentrate slurry is usually sent to a "ripening" tank which contains larger crystals. The small crystals melt because of their excess surface energy. This subcools the concentrate with respect to the large crystals, which consequently grow larger. Ripening produces large crystals, which can be cleanly separated from concentrate. The ice-concentrate mixture is often pumped upward through a wash column. Downward flows of water, obtained by melting the ice, displace and expel the concentrate from the upward-moving bed of large crystals. When low displacement velocities are used, very clean separations are obtained.

In theory, very little energy is required for freeze concentration. Freezing and ripening are carried out at a sequence of temperatures in multistage units, and discharged ice is used as a heat sink to provide higher energy efficiency and faster processing at lower cost. Nevertheless, energy requirements for freeze concentration are still higher than for evaporation.

Concentrations obtained by freeze concentration had been limited to 40–45%, because cold viscous concentrates could not be readily separated from ice. However, centrifuges have been developed to carry out this separation, and concentrations in the 50–60% range can be obtained. This is close to the theoretical maximum concentration (roughly 65% of sugar solutions) obtainable by freezing.

Reverse osmosis and ultrafiltration. In reverse-osmosis systems, water is forced through selectively permeable membranes by using pressures which exceed the osmotic pressure of the solution being concentrated. Reverse osmosis is primarily used to purify water and to concentrate milk, whey, and dilute extracts. Apple juice and maple sap are sometimes preconcentrated by reverse osmosis prior to evaporative concentration. *See* OSMOSIS.

In ultrafiltration systems, relatively low pressures are used to force water through porous membranes which retain high-molecular-weight solutes. Any low-molecular-weight solutes which are present pass through with the water. *See* ULTRAFILTRATION.

Reverse osmosis and ultrafiltration require very little energy, but cannot provide food solute concentrations which are much higher than 20%. Freeze concentration is much more expensive than evaporation. Because of these limitations, evaporation will continue to be used to concentrate liquid foods in most cases. Henry G. Schwartzberg

Dehydration

Drying is one of the oldest methods of food preservation; it is also one of the most difficult to model and control. Dehydration is usually considered a supplement of drying. A substance can be dried beyond its capability of restoration, but a dehydrated product is essentially in its original state after adding water. The distinction between evaporation and drying is subtle. Evaporation generally refers to removing water vapor at its boiling point. In drying, the water is usually removed below the boiling point as vapor by air. Many techniques have been developed to dry food products, but the mechanisms involved in the process are always the same. The basis of dehydration is simple: drying transfers energy into the product and transfers moisture out. The energy supplies the necessary heat, usually the latent heat of vaporization, and is responsible for the water migration. Materials may gain or lose moisture, depending on the relative humidity and temperature of the surrounding environment.

The purpose of food dehydration is to decrease microbial growth, enzyme activity, and deteriorative reactions by removing the available water in the product. Many foods are dried to a final moisture content of about 5%, but this varies with the food, package, and intended use. Dehydrated foods are ideally suited for prolonged storage, and among other benefits, dehydration may enhance the texture and flavor of the product or preserve desirable nutrients. Dehydrated foods lower transportation costs due to the reduction in product volume, and thus are conveniently handled.

However, consumer reluctance to accept the product can be a major drawback of dehydrated foods. Drying typically results in a number of changes to the material. These include bulk density change; case hardening and toughening; heat damage, which may include browning; effects on enzyme activity, microorganisms, pigments, proteins, fats, and carbohydrates; loss of ability to rehydrate; loss of nutritive value; loss of volatiles, for example, flavor; migration of soluble constituents; and shrinkage. These changes may or may not be desirable.

Successful design of a drying process requires a complete knowledge of the changes in the product during dehydration and an understanding of the variables which may affect the drying rate. This entails detailed information of the type of food being dried and its physical properties. Water availability and the type of so-called bound water present, as well as drying-rate concepts, are other important parameters to consider.

Water availability. Dehydration is primarily a preservation technique. Dehydrated foods remain shelf-stable without sterilization due to the small amounts of available water in the product. Microorganisms, enzymes, and other biological and chemical reactions cannot function in the absence of water. The availability of water is determined by the moisture content, water molecule interactions, and food component structure. Water binding is a measure of the molecular interactions and may be estimated by various methods. The most descriptive of these determines water sorption isotherms (**Fig. 25**), a relationship between the moisture content in food and the water activity. Water availability, then, rather than moisture content, is the influencing factor when dehydrating a food product.

Water activity. The water activity a_w is defined in terms of the mole fraction of the solvent, since at equilibrium the mole fraction of water in food is usually unknown and not easily measured. In addition, most foods do not behave as ideal solutions because of capillary forces, chemically bound water, and incomplete ion dissociation. Since a_w is equal to the water concentration in the liquid phase, it is the ratio of the water vapor pressure of the solution to the vapor pressure of the pure water at the same temperature, and therefore the nonideal behavior of food will be accounted for. Water activity indicates the relative ease with which water could be removed from a product during drying, since a_w expresses the relative humidity at which products are in equilibrium with their environment. Lowering a_w generally lowers the susceptibility of food to microbial spoilage. Typically, a value $a_w < 0.90$ prevents bacteria growth, $a_w < 0.80$ inhibits enzyme activity, $a_w < 0.60$ inhibits Maillard reactions, while $a_w \approx 0.20$ is the value needed for dehydrated food. *See* VAPOR PRESSURE.

Bound water. Bound water may exist under several conditions. Unbound water is found in cell pores or fibers, where soluble constituents cause a lowering of vapor pressure. Capillary forces also tend to exert a lowered vapor pressure on free water due to the concave curvature of its surface.

In food material, three types of water are usually present. These are as follows, in order of decreasing difficulty to remove: water bound to ionic groups, water molecules that are hydrogen-bonded, and

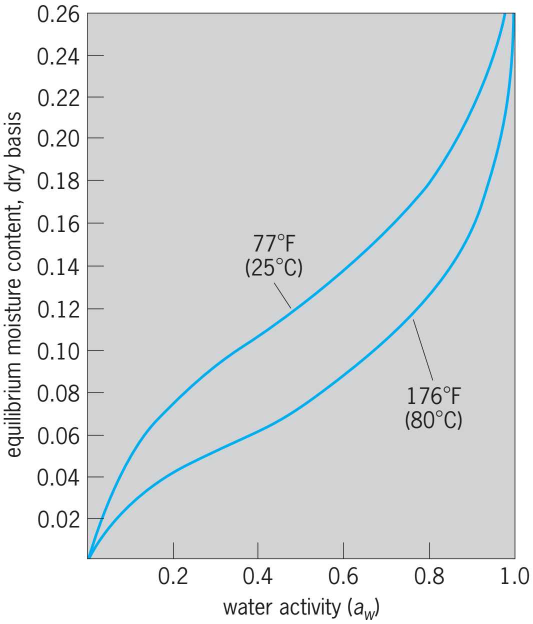


Fig. 25. Equilibrium isotherms for wheat. The values for dry basis represent unit mass of water per unit mass of dry solid. (*After M. Fortes, M. Okos, and J. Barret, Heat and mass transfer analysis of intrakernel wheat drying and rewetting, J. Agr. Eng. Res., 26:109–125, 1981*)

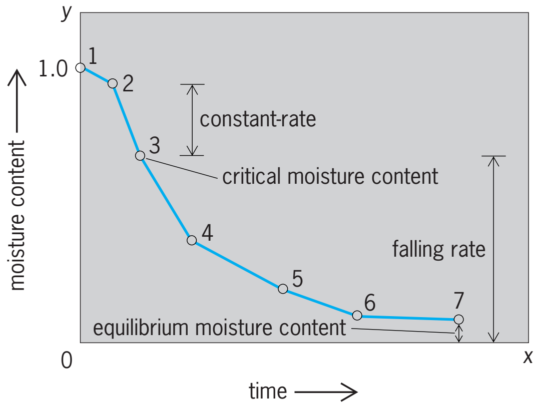


Fig. 26. Typical drying-rate curve for a biological material.

unbound free water. Certain energy requirements are associated with each level of water removal: first free water is removed, followed by hydrogen-bonded molecules, and then ionic group attachments. Since only a certain amount of energy is available in drying operations, different types of bound water are removed at different rates.

Drying rates. Criteria that hold for most dehydration processes are the characteristic drying regions of a material, which are known as the constant-rate period, the transition zone, and the falling-rate period. Each region is characterized by the type of resistance that controls the process. In the constant-rate region, external resistances control mass transfer. In the falling-rate period, internal resistances control mass transfer. The transition is the interface of the two, and is controlled by both mechanisms. It also contains the critical moisture content, defined as the moisture content where the internal resistance equals the external resistance.

In the constant-rate period a nearly linear relationship of moisture loss over time exists (**Fig. 26**). The rate of moisture removal is limited only by the rate of evaporation, and is dependent on air velocity, temperature, humidity, piece size, and bed depth. In this period, unbound, free moisture is always present on the surface of the material being dried. Moisture is able to diffuse to the surface faster than the rate at which it leaves, and the rate of evaporation is independent of the material. The drying rate is usually determined by the heat transfer. If heat addition is predominantly by forced convection, the surface temperature can usually be approximated as the wet-bulb temperature of the air stream. If the heat transfer occurs entirely by conduction, the surface temperature of the material is approximately equal to the boiling point temperature. If both conduction and convection heat transfer takes place, the surface temperature of the material will be somewhere between the wet-bulb and boiling point temperature. If heat transfer is by radiation or dielectric heating, other more complex relations must be used.

At the point of critical moisture content, the drying rate begins to proceed at a decreasing rate. After that point, an equilibrium moisture content is approached. This is the falling-rate period (Fig. 26), and there may be more than one. The drying rate is usually affected more by changes in air temperature and humidity and by piece size than by other conditions. The water within the material can no longer diffuse to the surface at a rate fast enough to keep the entire surface saturated with moisture, and the surface area decreases. The drying rate subsequently decreases as the unsaturated surface area increases. The material properties become very important and influence the drying rate. This is because moisture movement in the falling-rate period is in part determined by capillary flow, liquid diffusion, surface diffusion, and water-vapor diffusion, and is a function of the material.

Modeling the constant-rate region can be accurately accomplished by application of heat- and mass-transfer equations. The falling-rate region, however, is much more difficult to assess theoretically because the mechanism of transport is difficult to determine, it may involve a variety of mechanisms, and it is affected by external factors such as shrinkage and case hardening and toughening.

Categorization. Drying occurs by simultaneous heat and mass transfer, and may be classified in several ways. When drying is categorized according to the condition of heat transfer to the product, two general processes are considered: direct and indirect heating (**Fig. 27**). In direct drying, heat is added by direct contact to the product. Heated air is circulated in a chamber and forced over the food. The moisture saturated air leaves and removes the water vapor of the product. In indirect drying, heat is transferred indirectly to the material through an intervening medium. The product may receive heat through a plate or tray on which it is placed, or by radiation. A vacuum is usually applied to remove the water vapor, and to decrease necessary temperatures of vaporization. For instance, indirect vacuum drying would be

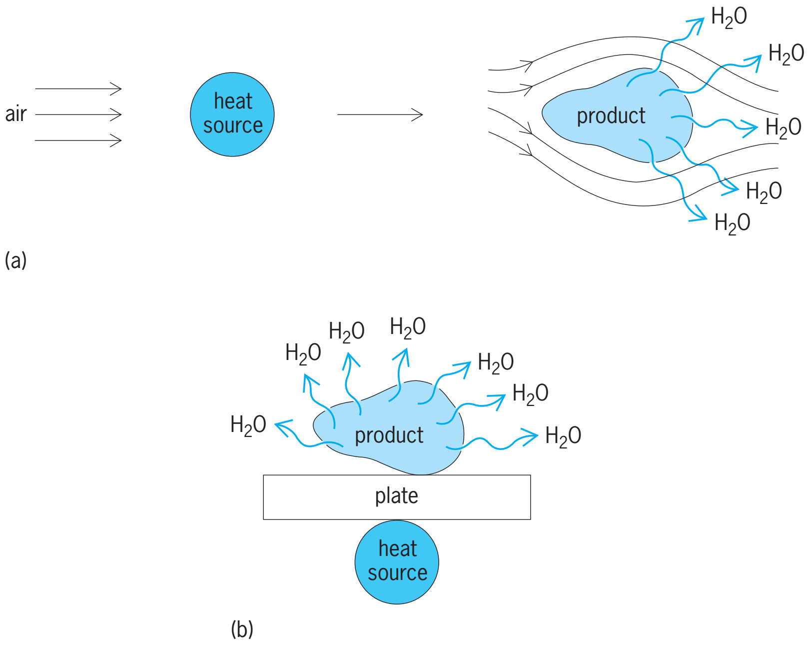


Fig. 27. Heat transfer to product by (*a*) direct and (*b*) indirect drying.

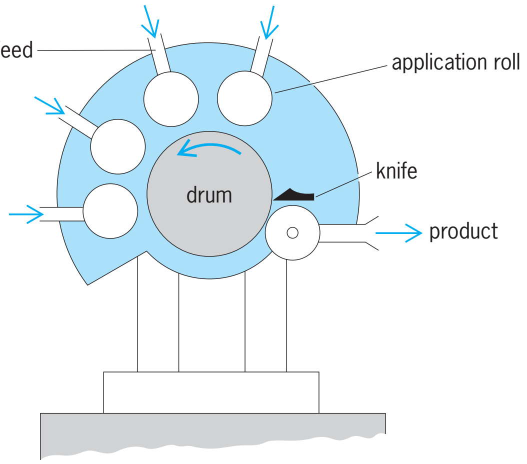


Fig. 28. Elements of a single-drum dryer.

used where a product is too sensitive to withstand exposure to a moving hot air stream.

Equipment and methods. A variety of equipment and methods is used in the drying process. The dryer chosen is dependent on product characteristics, desired quality retention, economic factors, and desired quality factors. Direct dryers include bed, belt, pneumatic, rotary, sheet, spray, through-circulation, tray, and tunnel. Indirect dryers include agitated-pan, cylinder, drum, freeze, microwave, screw-conveyor, steam-tube, vacuum, and vibrating-tray. Drum drying and spray drying are discussed below as important examples of typical drying methods.

Drum drying. Drum dryers utilize heat from the condensation of steam vapor on the inside of a drum surface to dry material that is contacted to the outside surface. Heat transfer is by conduction, and provides the necessary latent heat of vaporization. The drum rotates, and a thin layer of product is applied. Drying rates depend on film thickness, drum speed, drum temperature, feed temperature, ambient conditions, and product characteristics. The speed of rotation is adjusted so that the desired moisture content of the scraped product is obtained. Quality reduction is minimized by assuring that the product film has uniform thickness and that the dry film is removed completely. Drum dryers are very efficient, with heat loss typically 10% of the total input. However, drum dryers are not used extensively in the drying of food products due to the difficulty of applying a thin layer to the drum, and due to the long contact time at high temperatures, which may cause undesirable quality changes to the food product. Little design information exists on the subject. Most of the theoretical approaches account only for the heat transfer limitation when modeling dryer performance, when in theory drum drying is mass-transfer-limited.

The most important advantage of drum dryers is their capability of handling slurries and pastes of high moisture and viscosity. Products like mashed potatoes and powdered milk are typical applications. The ability of drum dryers to handle these difficult products makes them nearly irreplaceable in some operations.

Single-drum (**Fig. 28**) and twin-drum dryers are the most common types. Vacuum drum dryers have a vacuum tank which surrounds the drying surfaces. This feature decreases the temperature of operation.

Spray drying. Spray drying is a method of producing a dried powder out of feed in the form of pastes or slurries, or other liquids containing dissolved solids. Spray drying can be considered a two-stage process consisting of atomization and moisture removal. Atomization breaks the feed into a spray of liquid droplets. This provides a large surface area for moisture evaporation. The spray is contacted with a hot gas, usually air, and evaporation and drying produce the desired solid product particles.

Spray drying is an important preservation method in the food industry, since it removes moisture quickly and continuously without causing heat degradation. Heat damage to the food is prevented, due in part to the rapid drying cycle, the short total residence time, and cooling effects of evaporation. The disadvantages common to spray drying are the high energy needed and the requirement that the product be capable of being atomized. Furthermore, the product should not adhere to the interior of the dryer.

Spray drying is classified by the direction of product flow. There are cocurrent, countercurrent, mixed-flow, and parallel-type spray dryers. Cocurrent flow is best suited for heat-sensitive products.

In a cocurrent open cycle (**Fig. 29**) feed is pumped from the feed tank to an atomizer, where it is sprayed as tiny droplets inside the drying chamber. Air from the ambient surroundings is drawn in by a fan, filtered, and heated in an air heater. The hot air is contacted with the atomized food in the vicinity of the atomizer. Transport through the drying zone takes a matter of seconds. The powder falls to the bottom of the chamber, where it is pneumatically transported to the packaging line. The remaining product is separated in a cyclone, or a series of cyclones. The fines are recovered and transported to the pneumatic conveying system, while the air is vented to a filter and then to the atmosphere.

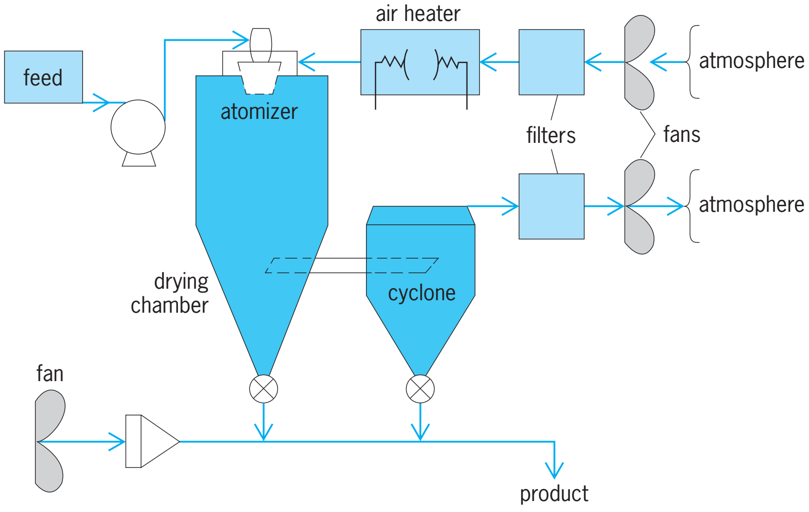


Fig. 29. Cocurrent open-cycle spray dryer layout.

There are many variations in spray dryer layout and design, depending on the requirements of the dried product. Additional equipment can be incorporated in the layout, and include wet scrubbers, fluidized-bed filters, and washing equipment. Variations in layout include open cycles as opposed to closed and mixed flow, and multiple-stage layouts that employ more than one spray dryer. A spray dryer may also use a direct-fired burner, utilizing the combustion of natural gas. Modifications of the spray dryer include a variety of atomizers (centrifugal pressure and fan-spray nozzles, two-fluid and rotary atomizers), and different shapes and sizes of the drying chamber. *See* DRYING.

Martin R. Okos

Expression

Expression is a process in which compaction caused by the action of a moving surface expels fluid from partially confined fluid-rich solids and solid-liquid mixtures, which are known as the press cake or cake. The process is used to recover oils and fats from oilseeds and rendered meat and fish scraps; reduce oil and water contents in fish, meat, bone, and poultry meals; dewater felts and mats formed from fibrous slurries; expel juice from fruits, sugarcane, and sorghum; expel pectin liquor from cooked pomace; dewater wet processing wastes; recover usable juices from such wastes; recover leaf protein; expel water or solvents from filter cakes; adjust the fat content of cocoa; dewater waste-treatment sludges; and expel whey from cheese.

The media. While most of the cake is confined, part is covered by a perforated surface, the media, or in special cases not covered at all. During compaction, expelled fluid flows out through the media or through the uncovered surface. The media often consists of light or heavy sheet metal with a uniformly spaced array of small holes, sometimes backed up by a fine-mesh screen. In some cases, a filter cloth supported by corrugated plates or grids is used as the media. When very high pressures are used, a cage made of sturdy, wedge-shaped bars separated by thin slits and backed by heavy ribs is used.

Equipment. Expression equipment includes screw presses (**Fig. 30**), in which compaction is induced by progressive reductions in the clearance between the root of the screw and the cage in which the screw is housed, or by reductions in the distance between screw flights or the clearance between adjacent twin screws. Compaction may also be induced by pressure developed by throttling the solid discharge at the end of the screw.

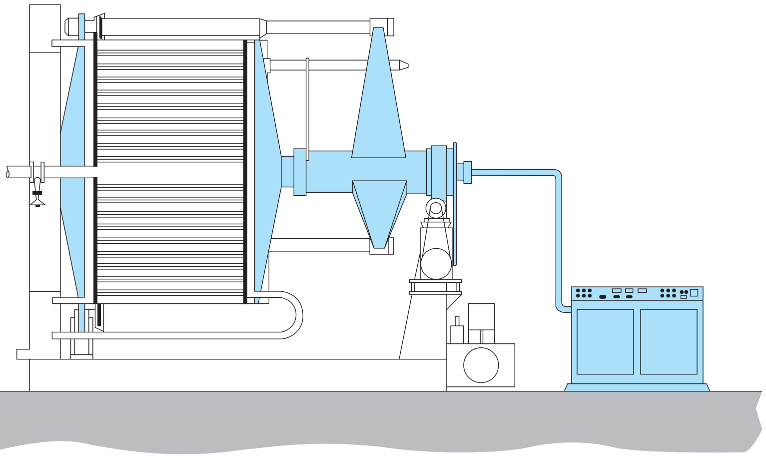


Fig. 31. Bucher-Guyer multitube press.

Cakes are also compacted inside perforated housings by solid or perforated rams or, in one case, by an inflatable rubber tube. In one widely used press, comminuted fruit is pumped into a cylinder that contains many cloth-covered, ribbed, flexible tubes attached to manifolds in the ram surface and the opposing stationary surface (**Fig. 31**). These tubes, which support the cloth media, convolute as the ram advances. Hence, the outflow area remains constant.

In rack-and-cloth presses, the cake is confined in flat pouches formed by folding over the corners of square filter cloths. These pouches are sandwiched between racks that provide mechanical support and clearance for juice drainage. Piles of the racks and pouches are compacted between platens driven together by screws or hydraulic pressure.

In belt presses, cake is deposited on a perforated moving belt and covered by a similar belt moving at the same speed. The belts are pressed together by passing them between a succession of rollers, or by placing them in tension and passing them over a rotating drum or drums. Juice yields and dewatering are improved by pulling belts over drums that provide S-like turns in the pressing path.

Sugarcane and sorghum are pressed by passing them through the nip between co-rotating, grooved rolls in three roll mills. The rolls are pressed together by hydraulic pistons. The expelled juice drains backward through solids advancing into the nip. In lighter

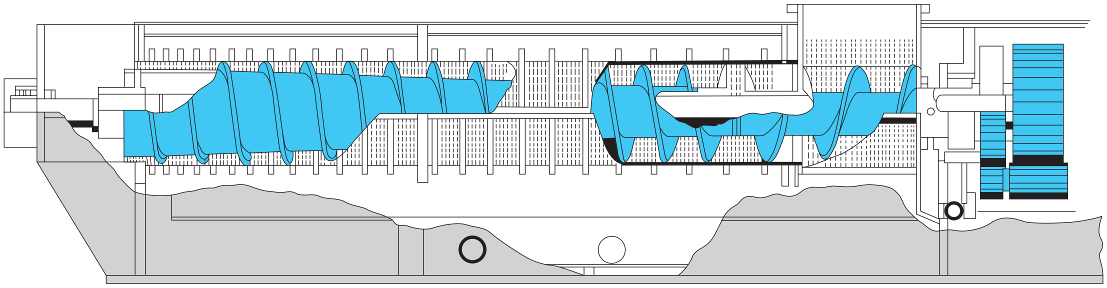


Fig. 30. Medium-weight screw press.

presses, which are used to expel juice from forage crops and dewater spent brewing grains, the solids pass through the nip between two co-rotating conical disks.

Solid-bowl, scroll centrifuges are sometimes used to press first-run grape juice from crushed, destemmed wine grapes and are often used to press olive oil and fruit water from milled olive paste. The scroll conveys solids toward the narrow, solids-discharge end of the centrifuge. Centrifugal force compacts the solids, forcing out fluid, which forms a pool over the compacted solids. The fluid discharges through holes at the wide end of the centrifuge. The process is very rapid, but compaction is not stringent. Consequently, yields tend to be low. Water is mixed with olive-paste feeds to improve yield by augmenting the volume of expellable fluid. This increases water-removal loads when expelled olive oil is subsequently separated from expelled fruit water.

Pressure and pressing rate. Pressing productivity and juice yield are governed by the relationships between the pressure used to achieve compaction, the specific volume of the presscake, and the rate of pressing. The pressure is the sum of (1) the stress developed in the solid itself or in the solid occluded-fluid ensemble; (2) compaction stress that is transferred from the presscake to adjacent surfaces by friction; and (3) the fluid pressure drop generated by fluid flow through the media and interparticle pores in the cake. The solid compaction stress depends mainly on the extent of compaction, and often increases exponentially as the cake volume decreases. It also increases, but to a lesser extent, as the rate of compaction increases. Friction reduces available compaction stress at a rate which is proportional to the compaction stress itself and the sliding-surface area.

Fluid pressure drop is proportional to the rate of pressing, the fluid viscosity, and the sum of the flow resistances of the media and cake. These resistances increase rapidly as cake volume decreases. Therefore, fluid pressure drop is often the largest compaction-pressure component at the end of pressing.

Methods have been developed for (1) measuring and correlating the forces and flow resistances encountered during expression; (2) predicting interactions between these forces and resistances and the cake specific volume; and (3) predicting juice yields as a function of pressing rate and maximum allowable pressure.

The allowable rate of pressing and the maximum allowable pressure are in turn determined by either strength limits of the expression equipment; power limits for the equipment drive train; pressures at which cake extrudes through media; pressures at which undesirable particle deformation occurs; desired extents of fluid removal; or desired uniformity of cake compaction and fluid content. Depending on which limiting factor applies, the maximum pressures used range from 2 to 5 lb/in.^{2} (14 to 35 kilopascals) for pressing soft cheeses, 6 to 12 lb/in.^{2} (41 to 83 kPa) for hard cheeses, 200 to 450 lb/in.^{2} (1.4 to 3.1 megapascals) for fruit juice expression, 500 to 1000 lb/in.^{2} (3.5 to 6.9 MPa) for dewatering spent sugarbeets, 1000 to 2000 lb/in.^{2} (6.9 to 14 MPa) for dewatering spent coffee grounds, and attain 6000 lb/in.^{2} (42 MPa) in sugarcane mills, and as much as 16,000 lb/in.^{2} (110 MPa) in some vegetable-oil screw presses.

Henry G. Schwartzberg

Filtration

Filtration processes are used for the removal of solids from a fluid (either a liquid or gas) by utilizing a physical barrier (the filter medium) containing openings or pores of the appropriate shape and size. The major force required is hydraulic pressure to pump the fluid through the filter medium. In the other mechanical separation processes, such as sedimentation and centrifugation, the force required is gravity or centrifugal force, respectively. In addition, they also require a significant difference in density between the two phases being separated, whereas filtration requires only a significant size difference.

Filtration processes cover a wide range of particle sizes (**Table 4**). The sieving or screening operations use fairly coarse barriers of perforated metal or woven screens. The screens may be vibrated or rotated to improve filtration rates (**Fig. 32**). On the other end of the spectrum is reverse osmosis (or hyperfiltration), a membrane separation process designed to filter out dissolved molecules in the ionic-size range. The middle range of solid-fluid separations involves removal of insoluble particles of 0.05 micrometer and larger, and is commonly referred to as filtration or microfiltration.

TABLE 4. Classification of filtration operations

Process	Particle size	Example
Screening, sieving	>100 μm	Protein precipitates (casein, curd particles)
Filtration	10–100 μm	Yeast, seedcoat in oilseed slurries
Microfiltration	0.05–10 μm	Bacteria, blood cells
Ultrafiltration	1–20 nm	Protein, soluble starch
Reverse osmosis	<1 nm	Salt, sugars

Equipment. Equipment for solid-liquid filtration can be classified in a number of ways: (1) whether the retained particles (the cake) or the filtrate (the clarified liquid) is the desired product; (2) batch or continuous operation; (3) driving force, that is, gravity-, pressure, or vacuum-activated; (4) mechanical arrangement of the filter medium, such as

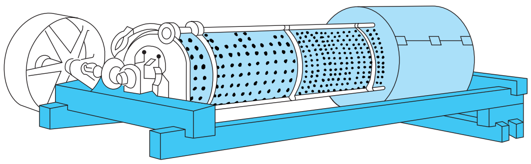


Fig. 32. Revolving screen filter. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, chap. 7, McGraw-Hill, 1956*)

plate-and-frame (**Fig. 33**), leaf, rotary drum (**Fig. 34**), or cartridge type; or (5) mechanism of filtration, that is, whether the separation occurs on the surface of the filter medium (as in membrane separations) or within the depths of the filter. Depth filters are usually used in dead-end-type equipment and in batch operations. An example is a clarification process in which the slurry has a low particle concentration, such as in the production of sparkling clear beverages. Surface filters are more suited for continuous operations where the retained particles have to be quantitatively recovered, such as the harvesting of microbial cells. *See* CLARIFICATION.

Filter media commonly used in industrial filtration include duckweave heavy cloth, paper, cellulose, metal cloth, nylon, and other synthetic cloths. Filter aids, such as diatomaceous earth or activated carbon, are also frequently used to improve filtration efficiency. Filter aids are applied as a precoat on the filter medium or added to the slurry before filtration. *See* ACTIVATED CARBON.

Theory. All models describing the filtration process must account for the buildup of solid particles on the filter medium and the resulting drop in filtration rate due to the increased resistance of the cake. Reasonably good mathematical approximations have been developed based on Poiseuille's law for laminar flow through channels, shown in differential form in Eq. (1), where V is the volume of filtrate in time t,

$$\frac{dV}{dt} = \frac{A^2P}{\mu\alpha CVP^s + AR_m} \tag{1}$$

A is the filter area, P is the pressure drop across the filter, C is the solids concentration in the slurry, μ is the fluid viscosity, α is the specific cake resistance, and s is a cake compressibility coefficient. The resistance due to the filter medium, R_m, is usually much less than the cake resistance and is usually neglected (that is, $R_m = 0$). *See* LAMINAR FLOW.

Filtration is conducted in one of two modes: constant-rate or constant-pressure. Most satisfactory results are obtained when a constant-rate operation at low pressure is used at the start of the process. Due to cake buildup, however, a gradual increase in pressure will be needed to maintain constant-rate conditions. When the pressure limit of the filter or the pump is reached, the operation is conducted under constant-pressure conditions until the filtration rate drops to an uneconomically low level. The process is then stopped to wash or remove the cake and clean the filter. The equation can be integrated with the appropriate boundary conditions to describe either the constant-rate or constant-pressure operation.

Optimizing filtration variables. The approximations based on Poiseuille's law in both its differential and integrated forms are useful in evaluating and optimizing operating variables, such as pressure, temperature, solids concentration, filtration rate, and filter area. Many of the variables involved are not mutually exclusive. It is almost always necessary to conduct experimental trials for every application in order to properly design a filtration operation.

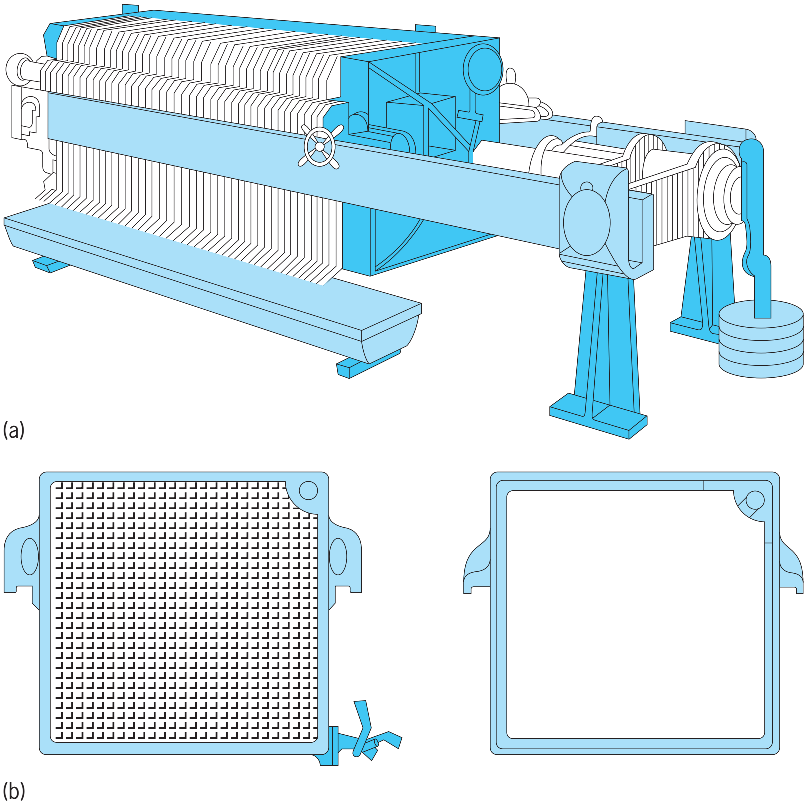


Fig. 33. Plate-and-frame filter press. (*a*) Outside view of press. (*b*) Plate and frame of open-delivery, nonwashing press. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, chap. 7, McGraw-Hill, 1956*)

Increasing pressure should increase filtration rate. This will be true for hard, granular cakes that are relatively incompressible (where $s = 0$ in the equation). On the other hand, if the cake is soft, gellike, and easily compressed (that is, where $s = 1$), the filtration rate will be independent of pressure, assuming $R_m = 0$. In practice, s has been found to vary between 0.1 and 0.8, and will have to be experimentally determined for each application.

Higher temperatures increase filtration rates, due primarily to its effect on lowering fluid viscosity.

Filtration rate can be increased by diluting the feed

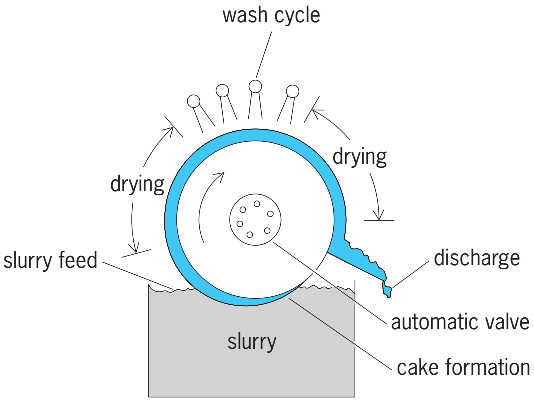


Fig. 34. Continuous rotary-drum filter. (*After C. J. Geankoplis, Transport Processes and Unit Operations, Allyn and Bacon, 1978*)

slurry. However, this must be balanced against the higher volume that must be processed.

The equation shows that the total volume filtered (V) is inversely proportional to the filtration rate (dV/dt). Higher volumes can be processed at lower filtration rates.

The filtration rate and the total volume filtered are proportional to the square of the filter area. Thus increasing filtration area may be the most cost-effective method of increasing filter throughout.

Membrane separations. Since the birth of modern membrane separations technology in 1960, a vast array of applications unmatched by any other processing technique in its variety have been developed. Desalination and water treatment by reverse osmosis is probably the earliest and best-known application. The 1970s saw increasing usage in the chemical process industries (paint, textiles, oil recovery, pulp, and paper). In the 1990s the most significant inroads of this technology were in the food and biotechnology processing industries, where membrane technology is finding increasing uses as a gentle and efficient way of fractionating, concentrating, and clarifying a variety of products.

Membrane separation processes are based on the ability of semipermeable membranes of the appropriate physical and chemical nature to discriminate between molecules primarily on the basis of size and, to a certain extent, on shape and chemical composition. The role of membrane is to act as a selective barrier, enriching certain components in a feed stream and depleting it of others. Membranes are made from a variety of polymers or inorganic materials; over 150 different materials have been used, with varying degrees of success. The food industry, however, uses only a few materials, primarily those that have met sanitation and regulatory standards. These include the cellulosics, ceramics, some polyamides,

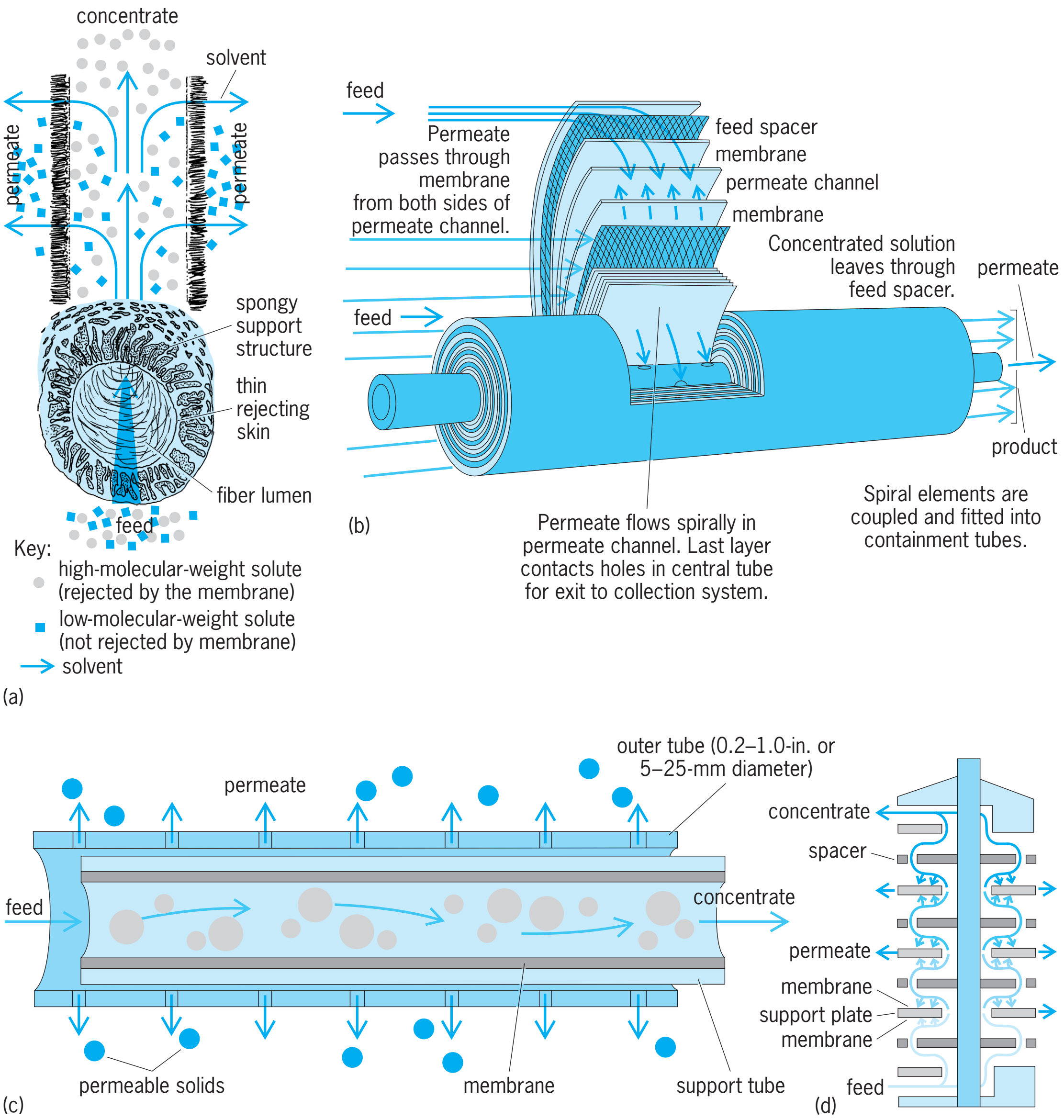


Fig. 35. Modules available for membrane filtration. (*a*) Hollow fibers. (*b*) Spiral-wound. (*c*) Tubular. (*d*) Plate system, showing internal flow.

polysulfone and its derivatives, polyvinylidene fluoride, and propylene. The chemical nature and physical properties of the membrane control which components are retained and which permeate through the membrane. For example, it is possible to concentrate or dewater by reverse osmosis, fractionate components in solution by ultrafiltration, and clarify slurries or remove suspended matter by microfiltration. Nanofiltration fits between reverse osmosis and ultrafiltration; nanofiltration membranes are designed to separate monovalent salts and undissociated organic acids from multivalent salts, sugars, and organic compounds. *See* CELLULOSE; CERAMICS; POLYMER.

One of the chief attractions of membrane technology is the low energy requirements compared to other dewatering processes. For example, open-pan evaporation may need over 0.27 kWh per pound of water removed (0.6 kWh/kg), multiple-effect or mechanical vapor recompression evaporation uses 0.020 kW/lb (0.045 kWh/kg), and reverse osmosis requires only 0.0023–0.0091 kWh/lb (0.005–0.02 kWh/kg).

There are several equipment types and module designs that, between them, will probably be able to handle any fluid stream that can be pumped (**Fig. 35**). The principal types are tubular, hollow-fiber, spiral-wound, and plate systems. Each design has its own special applications, advantages, and disadvantages (**Table 5**).

In selecting a particular membrane module, the major criteria are (1) the physical properties of the feed stream and retentate, especially viscosity; (2) particle size of suspended matter in the feed; and (3) fouling potential of the feed stream. For example, most food materials containing macromolecules display non-newtonian behavior. Their viscosity will increase dramatically above certain concentrations, making pumping difficult and reducing the mass transfer within the boundary layer. This will necessitate the use of modules that can withstand high pressure drops, eliminating most hollow-fiber and capillary modules. However, these modules have extremely high packing densities (surface area:volume ratios) and comparatively low energy consumption.

Similarly, large-bore tubular modules are more suitable for feed streams with large amounts of suspended matter such as citrus juices and animal waste streams, even though they have the lowest packing densities and highest energy consumption among all designs. Modules utilizing flat sheets (spiral-wound units, plate units, and pleated-sheet cartridges) usually have a meshlike spacer between sheets of membrane. This restricts their use to clear liquids or feed streams containing only finely suspended matter. However, spiral-would units are the lowest in capital costs and lowest in energy consumption, although this is somewhat offset by their generally lower flux with polarizing feed streams.

If there has been one single factor that has inhibited the large-scale use of the membrane technology, it is fouling of membranes. Fouling, that is, accumulation of deposits, manifests itself as a decline in performance, usually a decline in flux under constant operating conditions and a possible change in the sieving properties of the membrane. This phenomenon is frequently confused with other flux-depressing phenomena, such as change in membrane properties with time due to deterioration, changes in feed solution properties, and concentration polarization.

Minimizing membrane fouling must include consideration of the chemical nature of the membrane as well as physicochemical properties of the feed stream. For example, hydrophilic membranes (such as the cellulosics and ceramics) foul much less with protein-containing feed streams than hydrophobic membranes (such as polysulfone or polypropylene). With fermentation broths, although microbial cells may be the largest single component, it is rarely the cells themselves that constitute the main cause of fouling, but the media components, except perhaps at very high cell concentrations ($>$100 oz/ft^3 or 100 g/liter).

The parameter known as the cloud point is the temperature at which a dissolved solid begins to separate from a solution, imparting a cloudy appearance to the liquid. Antifoams, especially silicone-based, are particularly notorious foulants, while vegetable oils are the most benign. Antifoams with inverted cloud points (that is, cases where higher temperatures increase cloud formation) can cause dramatic reduction in flux, especially with hydrophobic membranes such as polysulfone, if operated above their cloud points.

Fouling can be minimized by (1) modifying the membrane chemically, for example, to make it more hydrophilic or to block certain functional groups that may be interacting with specific feed components; (2) by modifying the feed stream, for example, adjusting pH changes the charge or conformation of

TABLE 5. Features of membrane equipment

Features	Tubular	Hollow fiber	Plate	Spiral-wound
Channel height	>10 mm*	<3 mm*	<2.5 mm*	<1.5 mm*
Spacers	No	No	Yes/no	Yes
Prefiltration	No	Yes	Yes/no	Yes
Flow	Turbulent	Laminar	Laminar	Turbulent
Flux	High	Intermediate	Intermediate	Low
Area/volume ratio	Low	High	Intermediate	High
Energy/flux ratio	High	Intermediate	Intermediate	Low

*1 mm = 0.04 in.

a protein, which may minimize protein-membrane interactions, or it changes the solubility of salts so that they do not interact with or precipitate on the membrane; and (3) by physical or mechanical means, which may involve periodic backwashing, backpulsing, vibrating or rotating or spinning the module, pulsatile flows, maintaining a uniform transmembrane pressure with cocurrent permeate flow in the module, or electrically enhancing the removal of the polarized layer from the membrane.

On a worldwide basis, the dairy industry is probably the largest user of membrane technology in food processing, followed by the fruit juice processors. Ultrafiltration is ideal for clarification of clear juices, where the goal is to replace the conventional holding, filtration, and decanting steps, and perhaps the final pasteurization step. Higher yields; improved juice quality due to removal of pectins, polyphenol oxidase, and tannin-protein complexes; and elimination of filter-aiding and fining agents are its features. Concentration by reverse osmosis to high levels (>42° Brix; a concentration measurement) has been commercialized but is not widespread.

Other food applications include production of vegetable protein concentrates and isolates from soy, sunflower seeds, cottonseeds, peanuts, and so forth; removal of glucose from and partial concentration of eggs prior to dehydration; processing of animal products such as blood and gelatin; sugar refining, including desalting of molasses, purification, and preconcentration; removal of yeast from beer or alcohol distilleries by microfiltration or ultrafiltration instead of centrifuges or filters; corn refining, including clarification of dextrose and starch purification and concentration; and vegetable oil refining, primarily for degumming, hexane recovery, and deacidification.

Membrane technology possesses the necessary characteristics to meet the rigorous demands of biotechnology. The major applications are separation and harvesting of enzymes and microbial cells; continuous high-performance bioreactors for enzymatic and microbial conversion processes; tissue culture systems utilizing plant and animal cells as biocatalysts or for production of monoclonal antibodies and production of high-purity water. *See* BIOTECHNOLOGY; CHEESE; FILTRATION; MILK. Munir Cheryan

Mixing and Agitation

Mixing is carried out to provide homogeneous compositions and physical properties for food ingredients and combinations of ingredients. Mixing and agitation are also used to facilitate heat transfer and mass transfer, create dispersions, produce emulsions and foams, suspend solids, facilitate reactions, promote coalescence, and produce textural and structural changes.

Homogeneity can be characterized in terms of the smallness of local deviations from the average value of properties or component concentrations, and in terms of the domain size in which such local deviations are less than some maximum tolerable value. With liquid mixtures, homogeneity can usually be obtained over domains which involve tens of thousands of molecules; for solid mixtures, homogeneity can be provided only over domains which involve thousands or tens of thousands of particles (10^{20}–10^{21} molecules). Artificially induced circulation or agitation is frequently used for carrying out mixing.

Many different mixing and agitation methods are used in order to produce the liquid, solid, doughlike, and foamy mixtures and dispersions which are encountered in food processing. In general, these methods involve repeatedly subdividing the mixture into discrete domains which are then transported relative to each other so that the domains repeatedly encounter and exchange matter with new neighboring domains.

Liquids. Liquid mixing is usually carried out by using rapidly turning screw propellers or radial-flow turbines to circulate batches of liquid in baffled tanks (**Fig. 36**). The baffles, which are usually one-twelfth or one-tenth of the tank diameter wide, prevent the liquid from moving en masse in a vortex. Off-center and side-entering impellers can also be used to prevent vortex formation. Propellers and turbines with pitched blades (**Fig. 37***a* and *b*) provide fairly smooth axial flows which can be advantageously used for particle suspension, blending nonviscous liquids, and improving heat transfer between such liquids and heating jackets or coils. Straight-blade

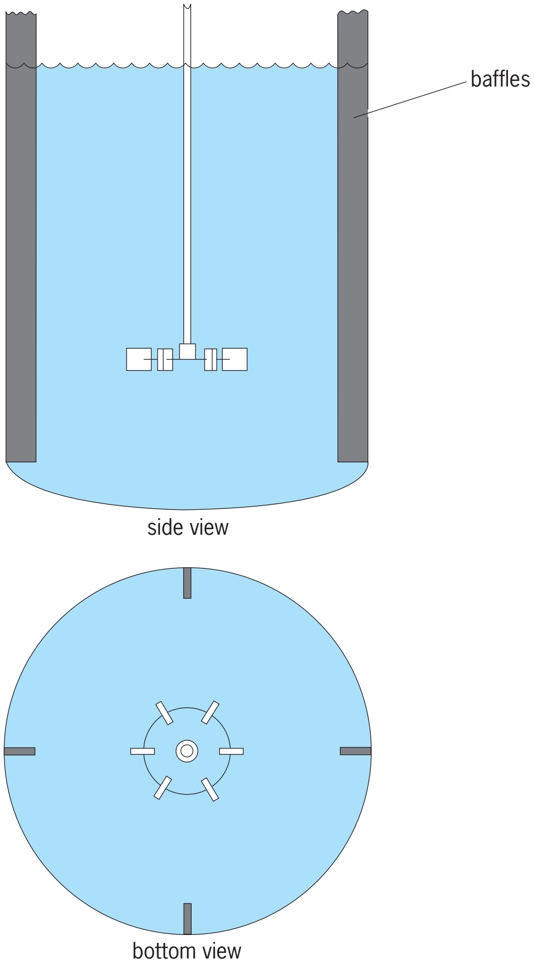


Fig. 36. Structure of a baffled mixing tank.

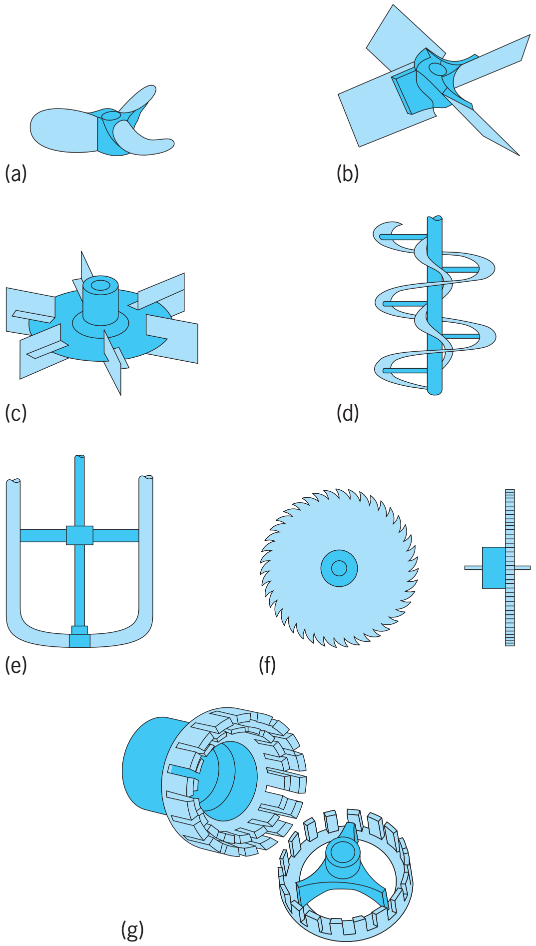


Fig. 37. Types of agitators. (*a*) Pitched-blade propeller. (*b*) Pitched-blade turbine. (*c*) Straight-blade turbine. (*d*) Double-helical impeller. (*e*) Anchor-blade impeller. (*f*) Sawlike impeller. (*g*) Slotted cagelike impeller.

turbines (Fig. 37*c*) are used to produce highly turbulent radial flows which are advantageous for producing dispersions and for facilitating mass transfer and heat transfer between suspended particles, drops, and bubbles and the liquid in which they are dispersed. Large-diameter, slow-speed, helical and double-helical impellers (Fig. 37*d*) and anchor-blade impellers (Fig. 37*e*) are used for mixing and circulating highly viscous and non-newtonian liquids. Sawlike impellers (Fig. 37*f*), slotted, cagelike impellers (Fig. 37*g*), and double counterrotating slotted cages are used to provide very high local shear rates and to create final dispersions. Various combinations of large-diameter, slow-speed helical mixers and high-speed, high-shear mixers are used to disperse and dissolve solutes which form viscous, non-newtonian solutions. *See* NON-NEWTONIAN FLUID.

The power P expended by impellers during turbulent vortex-free agitation of newtonian fluids is usually proportional to the value of the expression $C\rho N^3D^5$, where ρ is the liquid density, N is the rotational speed, and D is the impeller diameter. The constant C depends on the impeller type and tank geometry; it equals 1.0 for marine propellers and 6.0 for radial-flow turbines. The volumetric flow rate generated by impellers is proportional to ND_i^3. Heat-transfer, mass-transfer, and power-expenditure correlations for stirred tanks are frequently expressed in terms of Re_i, the impeller Reynolds number: $\mathrm{Re}_i = \rho ND^2\mu$, where μ is the viscosity. Heat-transfer coefficients for immersed surfaces in agitated tanks are usually roughly proportional to $\mathrm{Re}_i^{2/3}$. For propellers and pitched-blade turbines, fully turbulent flow occurs when Re_i is greater than 2000; for radial-flow turbines when Re_i is greater than 10,000; and for anchor-blade impellers when Re_i is greater than 100,000. Viscous flow occurs when $\mathrm{Re}_i < 10$. In the viscous range, power expenditures are proportional to μN^2D^3. *See* REYNOLDS NUMBER.

In the turbulent range, local mean velocities and fluctuating velocities in stirred tanks are proportional to $(PD/V\rho)^{1/3}$, where V is the tank volume. Mass-transfer coefficients for suspended particles, droplets, and bubbles, and the sizes of dispersed drops and bubbles are often correlated in terms of Reynolds numbers and Weber numbers based on the turbulent fluctuating velocity. Therefore, the ratio of power to volume (P/V) is often an important parameter in liquid mixing operations.

Recommended P/V values in horsepower per 1000 gallons (1 hp/1000 gallons = 0.2 kW/m^3) for typical liquid-food mixing operations are 0.25–0.33 for liquid blending; 0.12–0.25 for maintaining emulsion uniformity (for example, for preventing cream from rising in milk); 0.75–1.5 for improving heat transfer from fluids to immersed surfaces; 1.0–1.5 for dispersing starter cultures for cheese and yogurt manufacture, mixing ice cream ingredients, and dispersing and dissolving readily soluble solids; 4–10 for suspending rapidly settling particles (for example, hydrogenation catalysts), dissolving or dispersing sparingly soluble solids, and improving mass transfer between suspended particles and surrounding liquid; and 10–12 for dispersing air during aerobic fermentations. Power expenditures for suspending solids and blending liquids can be minimized by using large-diameter, slowly rotating propellers; but impeller and drive-train costs will be large when large-diameter propellers and slow rotational speeds are used. Therefore, small propellers and high rotational speeds are often used for small-scale or intermittent mixing.

Liquid mixing is usually carried out in tanks; in-line mixers, that is, agitators mounted in pipes, are also used. In-line mixers can be positioned to provide controlled holdup between mixing and subsequent processing, for example, between admixing a flocculant and separating resulting flocs. In-line mixers are also used immediately before equipment that is processing feeds that tend to separate, for example, to mix oil and aqueous liquid before entering a homogenizer. Static elements mounted in series in tubes can be used to provide in-line mixing. Elements split streams flowing through them and shift the positions of the splits. Adjacent elements create different splits and shift them in different senses. A series of staggered helical elements are often used. Successive elements turn in opposite directions and split the

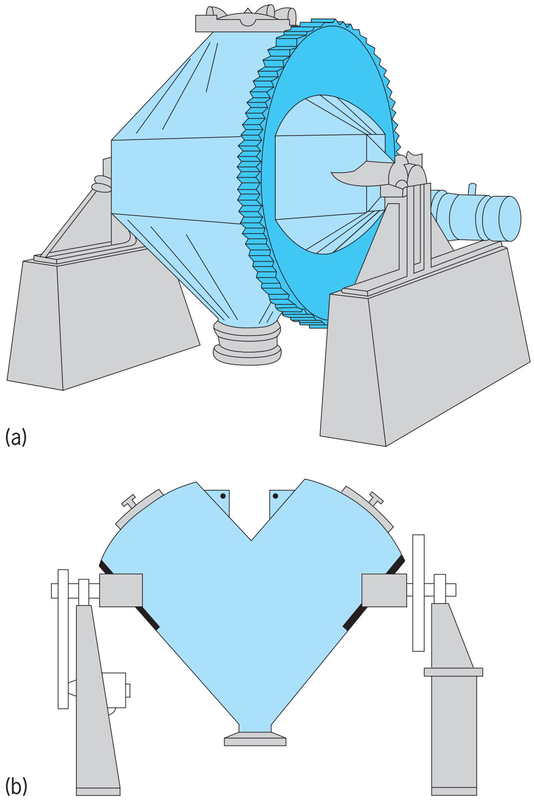


Fig. 38. Dry-powder mixers. (*a*) Double-cone blender. (*b*) Twin-shell (vee) blender.

stream 90° to the preceding split (**Fig. 38**). Mixtures split 2^n times in passing through n elements. Different flow-splitting arrangements are used in other static mixers. Static mixers are used to blend materials with widely differing viscosities, for example, frozen orange juice concentrate and water; blend shear-sensitive ingredients, for example, to mix blueberries with yogurt; carry out limited mixing, for example, to marbellize ice cream; and facilitate in-line heat exchange with viscous materials, for example, peanut butter, chocolate, and edible oils.

Solids. The mixing of solid powders is usually carried out in equipment which causes the solid mixture to repeatedly split into two or more portions, which are then recombined. In tumbling-type mixers (Fig. 38), mixing is brought about by lifting the mixture and letting it split as it falls in different increments as the mixer turns. Conveyors are used to transport part of the mixture relative to adjacent parts in ribbon blenders (where oppositely pitched, concentric pairs of ribbonlike, horizontal helices are used and in mixers where tilted, vertical screw conveyors which orbit around a conical hopper are used. Continuous, in-line mixing of solids can be provided by using twin-shell (vee) blenders mounted in series on an inclined shaft; screw conveyors with cut, cut-and-folded, or ribbonlike flights; or static mixers mounted in vertical ducts. Feed streams have to be automatically proportioned to obtain consistent uniform mixes with in-line units.

Solid mixing is characterized by diffusive processes, which tend to cause uniform spreading of the components of a mixture, and segregation processes, which tend to cause selective migration of certain components. Segregation is enhanced when the particle sizes and, to a lesser extent, the densities of components of mixtures differ. Small particles will frequently percolate through a bed of mixed particles and accumulate near the bottom of the bed, or in ribbon blenders they may be thrown into the air and temporarily adhere to the cover of the blender. In the latter case, erratic variations in mixing uniformity will occur.

Solid mixing usually results in random composition variations in local sample domains. The theoretical minimum standard deviation for the composition of a sample will be inversely proportional to the square root of the number of particles in the sample. Mixing indices based on the ratio of the theoretical minimum standard deviation to experimentally measured standard deviations are sometimes used to characterize the thoroughness of mixing and to correlate mixing rates. Maximum values for such mixing indices will frequently be in the 0.6–0.7 range—that is, substantially less than 1.0, particularly when segregation tends to occur—and are reached only after mixing times of 30 min to 1 h. By contrast, for comparable batches of liquids, very uniform mixing is usually achieved in 5–15 min. If the fine particles in a solid mixture tend to adhere to the larger particles, a more uniform type of mixing known as ordered mixing may be obtained.

The mixing of pastes and doughs is carried out in units in which the mixture is sheared and stretched between sturdy, slowly rotating elements with small interelement clearances. This process stretches and folds portions of the mixture and forces these sheared portions into other portions of the mixture. Relatively long mixing times and large power expenditures are required when doughs and pastes are mixed. Screw-based extruders are sometimes used for mixing doughlike products. Forward flow and transverse flow are induced by the shearing motion of the screw relative to the barrel, and reverse flow is induced by the pressure built up across the die through which the extruded product discharges. This combination of flows produces very intense mixing in the extruder.

The thoroughness and completeness of dough mixing can be characterized in terms of mixing indices, which are similar to those used for mixing powders. The nature and functional character of dough components—for example, gluten in wheat—and the overall rheological nature of a mixture often greatly change during mixing. In chocolate conchers, small particles of sugar are progressively broken down into very fine particles and coated with semifluid cocoa butter to produce a very smooth mixture. Textural and functional effects produced during mixing of doughs and pastes cannot be adequately measured or characterized by uniformity-based mixing indices. Power expenditure and agitator torque rise as bread doughs are mixed; the dough is considered fully developed when the torque reaches a maximum, and degrades with further mixing. The values of P/V for dough mixing are usually 33–134 hp/1000 gallons. *See* COCOA POWDER AND CHOCOLATE. Henry G. Schwartzberg

Extraction

Extraction is one of a family of separation processes employed for selective removal of components of a mixture by contacting it with an immiscible or partially miscible solvent to form a two-phase liquid-solid or liquid-liquid system (**Table 6**). Partitioning of solute molecules between the two phases provides the driving force for extraction. Based on the difference in either the solute's affinity for the solvent or its concentration in the two phases, the partitioning is conveniently expressed as a partition coefficient (K), defined as the ratio of a solute's concentration in the lighter phase to that in the heavier phase. Analytical data on the composition of immiscible phases in equilibrium can be used to determine K. The effectiveness of separating competing solutes is indicated by a separation factor (selectivity), which is the ratio of individual solute partition coefficients. Effective separation of one solute from another requires a separation factor significantly different from unity.

The most common solvents for extraction of food components are water, organic solvents such as hexane, and supercritical fluids. A fundamental understanding of rate-controlling steps governing the extraction of the solute of interest is needed in the design of an efficient process. Like other food processing unit operations, extraction can be carried out as a batch (unsteady-state) or a continuous (steady-state) operation. While a continuous system is more efficient in a large-scale operation, batch separation operations are common for biotechnologically produced products. In continuous, large-scale systems, the two phases flow either countercurrently or cocurrently, the former being more common. When recovery of the desired component is inadequate in one extraction, the process is repeated in a cascade of separation steps, and the process is known as staged extraction. The mixing and setting time needed in a stage extraction is eliminated by differential extraction in which the two phases continuously flow past each other without attaining equilibrium.

Liquid-liquid extraction. In this type of extraction, the solution containing a desired solute and the solvent, a liquid having preferential solubility for the solute, are mixed thoroughly. The solvent acts, and an immiscible layer, called the extract, is formed. The extract is separated from the remainder of the original solution, which, along with the residual solvent, is known as the raffinate. The solute in the extract may be purified in a secondary operation by using such techniques as membrane separation, electrophoresis, crystallization, precipitation, or distillation; and the solvent is reused after purification. Liquid-liquid extraction is ideal for recovering relatively nonvolatile and heat-labile solutes. Examples include vitamins, alkaloids, and pigments from animal or vegetable extracts. *See* CRYSTALLIZATION; DISTILLATION; ELECTROPHORESIS; MEMBRANE SEPARATIONS; PRECIPITATION (CHEMISTRY).

Solid-liquid extraction. Also known as leaching, solid-liquid extraction involves separation of soluble components from a solid by dissolving in a suitable liquid solvent. The size and shape of the solid matrix affects the performance of the extraction process. Increasing the surface-to-volume ratio through grinding, cutting, agitating, and increasing the temperature is effective in enhancing the rate of solute removal. If necessary, the solids may also be conditioned to release the desired component to the solvent phase. Often, solid-liquid extraction is aimed at recovering not one particular solute but all the soluble compounds in the solid matrix. For example, in the extraction of roasted coffee beans with water the extract contains all the water-soluble coffee constituents. *See* SOLVENT; SOLVENT EXTRACTION.

Supercritical fluid extraction (SFE). A supercritical fluid is any fluid at temperature and pressure above its critical values which, for example, for carbon dioxide are 31°C (88°F) and 72.9 atm (7.4 megapascals). In its supercritical state, a fluid exhibits physicochemical properties intermediate between those of liquids and gases, enhancing its role as a solvent. Its reversible high density gives good solubilizing power, while its low viscosity and diffusivity values provide appreciable penetrating and mixing power. These unusual properties have been found useful in extracting otherwise hard-to-separate constituents from their matrices.

Supercritical fluid extraction has emerged as a unit operation that exploits the unique properties of solvents to extract soluble components from a mixture. Partly because of the increasing regulation of solvents used in conventional extraction and partly

TABLE 6. Selected examples of extraction processes

Extraction	Material	Solvent	Solute
Liquid-liquid	Vegetable extract	Methylene chloride	Steroid
	Fermentation broth	Ethyl acetate	Antibiotic
	Cottonseed oil	Propane	Oleic acid
Liquid-solid	Ground roasted coffee	Water	Soluble coffee constituents
	Pressed olive pomace	Hexane	Olive oil
	Sugarbeets	Water	Sucrose
Supercritical fluid	Green coffee beans	Carbon dioxide	Caffeine
	Milkfat	Carbon dioxide	Cholesterol
	Lemon peel	Carbon dioxide	Lemon oil
	Hops	Carbon dioxide	Alpha acids

because of the potential advantages of better product quality, supercritical fluid extraction has found application in several industries.

Although a large number of supercritical fluids are suitable for food applications, supercritical carbon dioxide has been the solvent of choice because of its safety and its moderate temperature and pressure requirements.

A system for supercritical fluid extraction consists of four basic components: a solvent compressor or pump, an extractor, a temperature/pressure-control system, and a separator or adsorber. Additionally, other equipment, including ancillary pumps, valves, back-pressure regulators, flow meters, and heater/coolers for temperature control of the fluid, are needed for proper operation of the process. *See* SUPERCRITICAL FLUID. Syed S. H. Rizvi

Extrusion

Extrusion has become a widely practiced food processing operation used for the production of snack foods, ready-to-eat breakfast cereals, soup and beverage bases, pregelatinized starches, baby foods, candy products, full-fat soy flour, and textured vegetable protein. In addition to human foods, most dry and semimoist pet foods are also produced on extruders. The ability of the food extruder to perform both cooking and forming operations within a single machine having high capacity and continuous-processing capability has been the principal reason for the increased interest in the process. Other popular aspects of the process include its energy efficiency, a capability to handle a wide variety of food ingredients, and an ability to produce many types of products while creating no undesirable waste streams.

Food extrusion began with the development of the macaroni extruder in 1935. The macaroni extruder has a blender to combine the raw ingredients consisting of semolina wheat flour and water, and a transfer auger to move the mixture to the flighted extrusion screw which thoroughly mixes the ingredients and forces them through a die to form the shaped product.

Types of extruders. There are two types of extruders commonly used in the food industry. The most common is the single-screw type having one flighted Archimedes screw rotating within a stationary barrel to accomplish the processing. The operation of the single-screw food extruder is governed by the speed of rotation of the screw and the design of the flights along its length. Screws with shallow flights impart greater shear and work, rapidly increasing the product's temperature and producing higher pressures behind the die. Such equipment is used to manufacture completely precooked products such as ready-to-eat cereals, snacks, and textured vegetable protein. Extruders with deeper-flighted screws provide a less severe processing environment, which is necessary for the forming of macaroni or chewing gum without appreciable heating of these products. Most food extrusion equipment is of the single-screw type.

The second extruder type has a twin-screw configuration (**Fig. 39**) which employs parallel

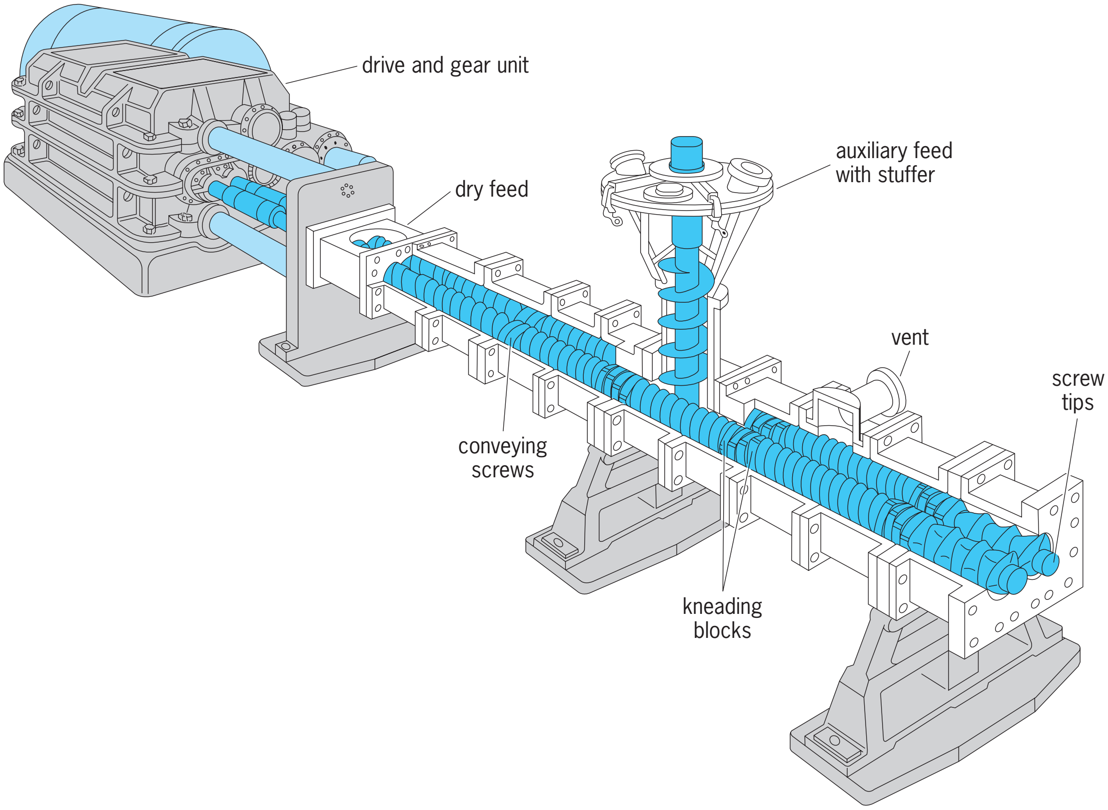


Fig. 39. Twin-screw extruder with cutaway exposing screw configuration. (*Werner and Pfleiderer Co., Ramsey, New Jersey*)

intermeshing screws which turn within a figure eight-shaped barrel opening. A typical twin-screw machine employs various screw elements for conveying, kneading (working), and pressurizing the sample. The advantages of the design are its greater ability to handle low-moisture ingredients, better control of processing conditions, and enhanced capability to convey ingredients.

Extruder operation. Major raw ingredients for extrusion processes consist of ground cereals or mill feed fractions, starch, vegetable protein meal, fat, and sugar. Minor ingredients include salt, acid or base compounds, emulsifiers, color, and flavor. Water is added to this mixture to achieve moisture contents ranging from 10 to 35% with a typical value of 25%. Ingredients may be preheated to almost 212°F (100°C) through the addition of steam to initiate cooking, which increases the capacity of the extruder and provides operating flexibility in a step called preconditioning. The preconditioner is a paddle blender capable of mixing liquid and solid food which is placed just before the extruder.

In a single-screw food extruder (**Fig. 40**) the granular raw or preconditioned ingredients fall from the feed hopper into the relatively deep pockets (channel) between the flights in the feed section of the screw. The flights on the turning screw continuously push the ingredients forward through the barrel and work them into a dough. Reducing flight heights down the screw's length decreases the area available for flow within the channels of the screw, which in turn partially restricts, the flow of the food ingredients during their passage. The restriction causes the channels of the screw to become completely filled with the food product, which increases the mechanical working of the dough. The mechanical energy input from the drive motor used to turn the screw is dissipated into heat by fluid friction and serves as the principal source of the energy necessary to cook the product. In addition to the mechanical energy heating the product, heat can be added through jackets surrounding the barrel and, in some cases, by direct injection of steam through the barrel wall. In this latter case, the action of the flights mixes all components within the screw channel to produce a relatively homogeneous mixture.

In cooking extrusion, the temperature of the food ingredients rapidly rises after they are compressed and worked to temperatures as high as 300–390°F (150–200°C) just before the die. Because the residence times in the extruder are less than 60 s, the process has been termed high-temperature short-time (HTST). High-temperature short-time processing preserves the nutritional quality of the food but inactivates enzymes, microorganisms, and other undesirable factors. An environment similar to that of a pressure cooker prevents boiling or flashing of moisture within the extruder because of the elevated pressures.

The most dramatic changes occur to the raw food ingredients in the compression section of the screw. The combination of heat and working converts the discrete particles in the raw feed into a continuous and very viscous dough. Physical changes accompanying this transformation include the hydration and disorganization of raw starch (gelatinization), the denaturing of protein including the chemical crosslinking of adjacent molecules to form a polymerized structure, and the melting of both starch and protein to give a translucent dough when the processing is complete. Typical doughs have much higher viscosities than bread dough because of the relatively low moistures involved in the process.

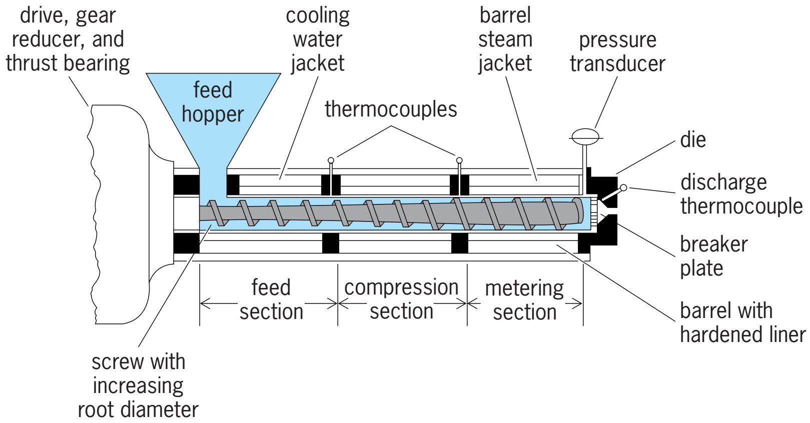


Fig. 40. Cross-sectional drawing of a single-screw food extruder. (*After J. M. Harper, Extrusion processing of foods, Food Technol., 32(7):67, 1978*)

The metering section of the screw has the shallowest flights, which enhances mixing to assure uniformity of the high-temperature dough. The metering section increases the pressure on the dough, which provides the force needed to push it through the die opening. The configuration of the die openings creates specific product shapes which are cut to a desired length with a rotating knife at the face of the die. Under high-temperature cooking extrusion conditions, puffing occurs once the dough exits the die and is exposed to atmospheric pressure. The puffing is the result of the expansion of the high-temperature water within the product. Loss of moisture accompanies puffing, causing the product's temperature to drop rapidly to about 140°F (60°C). Further cooling and drying are necessary to prevent adjacent pieces from sticking. Extruded ready-to-eat cereals and snacks exhibit an open cellular structure typical of extrusion-cooked and puffed cereal products. Protein flours extruded in a similar manner will puff, creating a chewy characteristic similar to meat because of the crosslinked protein molecules.

Extrusion processes. The extruder is just one of a series of pieces of equipment necessary to produce convenience food products. The sequence starts with bulk tanks that hold the raw ingredients in flour form (**Fig. 41**). These are metered into the extruder which cooks and forms them. The individual, cooked and expanded pieces are pneumatically conveyed to a dryer-cooler to reduce their moisture to less than 10%. Finally, flavor, oil, sugar, color, or vitamins and minerals are enrobed (coated) onto the surface of the pieces. To complete the process, the finished

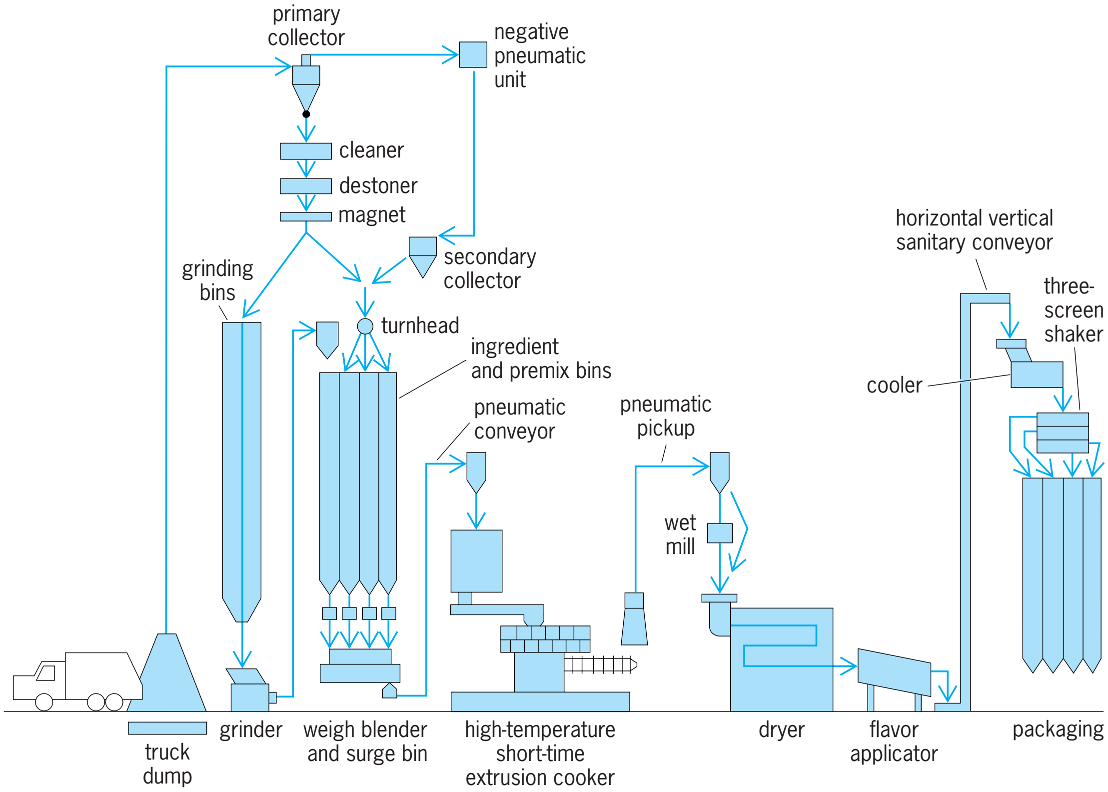


Fig. 41. Flow diagram for an extruded food process. (*After O. B. Smith and S. C. Crocco, Engineering "meat," Food Eng., 47(10):EF22, 1975*)

products are packaged in bags, boxes, or cans for distribution and consumption. Judson M. Harper

Cleaning and Sanitation

Generally speaking, the words cleaning and sanitation as used in the food industry have a similar meaning in that cleaning is recognized as being necessary for proper sanitation. Cleaning is the treatment given all product contact surfaces following each period of use in order to remove all physical evidence of residue. Typically, after cleaning is completed, it is followed by a chemical or thermal treatment to sanitize the product surfaces. In comparison, sanitation may be defined as the systematic control of environmental conditions during processing, storage, and transportation of food products to prevent contamination by animals, insects, and microorganisms.

Since the sanitation aspects of processing, storage, and transportation are specialty fields, involving many areas of expertise, this discussion will deal only with cleaning as related to processing operations in a food plant, with emphasis on clean-in-place methods.

Cleaning methods. The cleaning of product contact surfaces in a food plant can be carried out through the use of one or more techniques described in the following.

(1) In manual or hand cleaning of equipment, the equipment is disassembled, rinsed with a hose, and washed, typically using a mild wash solution and a hand brush.

(2) In semiautomatic cleaning, the equipment again is disassembled, rinsed with a hose, and then put in a clean-out-of-place tank (**Fig. 42**), which provides mechanical action of the wash solution through the utilization of a recirculation pump, and chemical action via the detergent solution which is maintained at a preset temperature.

(3) Spray or foam cleaning, utilizing spray wands, is frequently employed in the food industry

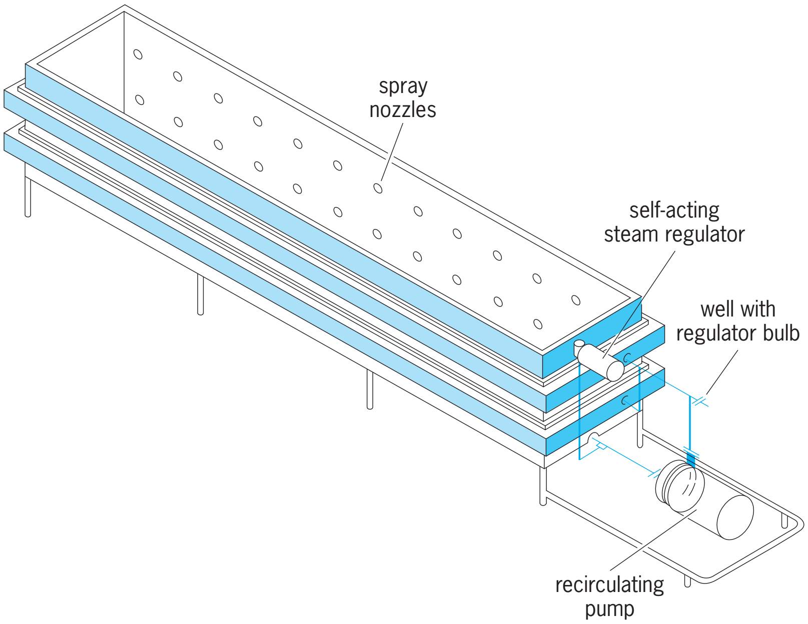


Fig. 42. Diagram of a clean-out-of-place cleaning system.

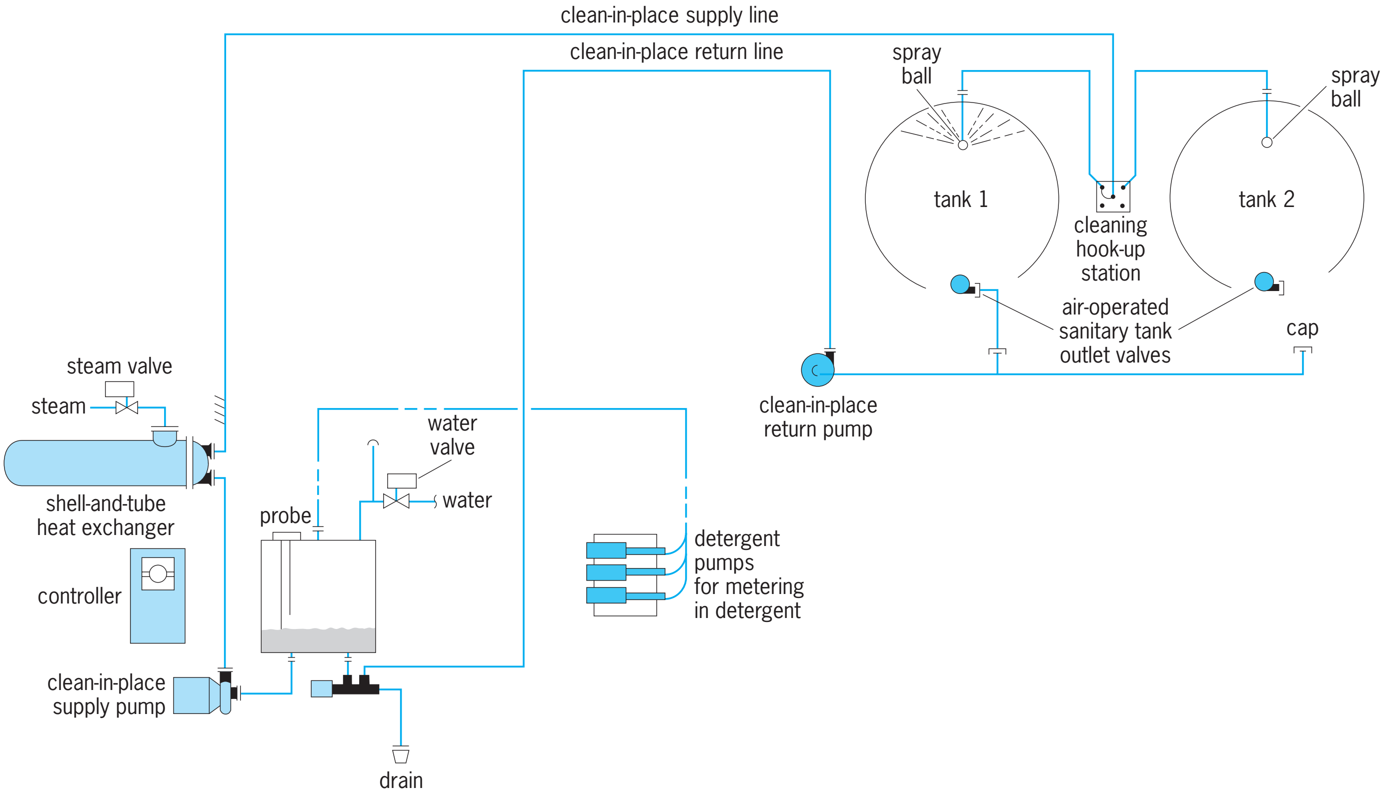


Fig. 43. Clean-in-place single-tank–single-use recirculating system.

(particularly in meat processing plants) to clean both the exterior and interior of equipment (which is not usually designed for automated cleaning) without any major disassembly of the equipment. Either low pressure (less than 100 lb/in.^{2}, gage, or 690 kPa) or high pressure (approximately 500 lb/in.^{2}, gage, or 3.5 MPa) sprays may be used for these operations.

(4) Clean-in-place is becoming the preferred method of cleaning the internal surfaces of processing equipment and the associated product pumps, sanitary automatic valves (and valve groups), and sanitary product transfer lines. The accepted definition of clean-in-place is cleaning entirely by circulation and flowing chemical detergent solutions and water rinses onto and over the surfaces that are to be cleaned by mechanical means. Cleaning-in-place techniques are utilized for cleaning both the inside surfaces of product transfer lines and the processing equipment. Generally speaking, the utilization of cleaning-in-place for product lines can be a relatively simple procedure if there are no product valves associated with the cleaning operation. However, with the incorporation of sanitary air-operated valves (such as valve groups), the proper directional flow of the clean-in-place solution through the valves can be quite complex. Cleaning-in-place for processing equipment such as storage tanks (horizontal or vertical), processors, pumps, blenders, heat exchangers (plate or tubular), evaporators, and dryers is generally complex, because not all of the equipment is specifically designed for clean-in-place applications and generally requires the addition of the proper type of spray device to assure complete coverage of all product contact surfaces, including agitator assemblies and baffles. The utilization of a permanently installed, nonrotating type of spray device of sanitary design is preferred. A simple unit is available that can be applied to cleaning operations for both product lines and processing equipment (**Fig. 43**).

Factors affecting cleaning. Regardless of the techniques utilized to clean product-contact surfaces of processing equipment, all of the food industry soils are composed of the same basic ingredients—protein, fats, carbohydrates, sugars, and mineral salts.

Soil removal through the use of detergent solutions is dependent upon four functions: (1) concentration of the detergent solution; (2) temperature of the detergent solution; (3) time that the detergent solution is in contact with the product surface; and (4) velocity or mechanical action of the detergent solution. As a rule, if the concentration of the detergent solution is increased, either the temperature of the solution or the time element can be reduced, with the same overall cleaning effect. This general relationship would also apply for increasing the temperature or the time to reduce the concentration of the detergent solution.

Detergent selection. Cleaning compounds are required to assure complete removal of all residues from product contact surfaces. Selection of a cleaning agent depends upon the composition of the soil, the nature of the soil, methods of cleaning, and materials of construction of the equipment being cleaned. Alkaline cleaners are effective in removing residues composed of fats, carbohydrates, sugars, and proteins. High levels of alkalinity (caustic) are required for effective removal of proteins. Sodium hypochlorite, when added to an alkaline cleaning solution, can aid in the removal of proteins at lower levels of alkalinity. Mineral salts, whether a component of the food product or residues from water, require acid cleaners.

Alkaline cleaners used in clearing-in-place are generally based upon sodium hydroxide (caustic) together with water conditioning agents, dispersants, and special low-foaming surfactants. Acids detergents used in cleaning-in-place are often based upon phosphoric or combinations of phosphoric and nitric acid with or without low-foaming surfactants. Heavy soils often require both an alkaline and acid detergent applied in two steps. Acid-based sanitizers can be used to remove light mineral films and at the same time sanitize the equipment.

Detergents used for cleaning-in-place can be highly active with an alkaline pH in excess of 11.5 or acids with pH less than 2.0. Temperatures of 140–180°F (60–80°C) are typical, depending upon the equipment designed and amount of soils present. For the cleaning of non-stainless-steel equipment, such as aluminum, special noncorrosive alkaline cleaners using silicates are required.

Following cleaning and rinsing, a treatment with a chemical sanitizing agent is required to assure a sanitary environment for the production of high-quality foods. In the United States, sanitizers must be registered with the Environmental Protection Agency (EPA) and comply with the regulations of the Food and Drug Administration (FDA) as an indirect food additive. Sanitizers used in the food processing industry are based upon chlorine, iodine (iodophors), quaternary ammonium chlorides, acids, and peracetic acid. Registered sanitizers, when applied as labeled to surfaces that have been cleaned, assure a level of sanitation that protects the safety of the food supply and meets quality concerns of the producer. *See* DETERGENT; SURFACTANT.

Equipment design. There are some guidelines that must be followed to ensure proper cleaning of the food processing equipment. (1) The processing equipment must be constructed of inert, nontoxic, nonporous material such as stainless steel; plastics and other synthetic materials must be approved by the FDA or the Department of Agriculture. (2) Processing equipment should be designed according to sanitary design specifications based on standards of the Technical Committee of the Dairy and Food Industries Supply Association. (3) Processing equipment should be specified to incorporate "3-A" sanitary standards, which assures that the equipment is properly designed from a sanitation standpoint and uses only properly approved contruction materials. (4) In addition, the processing equipment, pumps, valves, and associated equipment should be designed for clean-in-place by the manufacturer. The design aspects necessary to meet this requirement can be complex, but they include such items as utilization of radius corners of $^1/_8$ in. (0.3 cm) or more rather than the more conventional 90° angle; smooth, self-draining surfaces; ease of disassembly; and ease of access to all product-contact surfaces for hands-on or visual inspection. The first attempts at in-place cleaning were carried out on the plate heat exchangers utilized in the dairy industry for high-temperature short-time pasteurization of milk. Initially, powdered detergents were manually added to the balance tank of the high-temperature short-time system, and the product pumps were used for recirculating the cleaning solution through the heat exchanger.

Early work involving in-place cleaning of storage vessels utilized specially designed spray balls which were inserted into the access hole of the storage tank to assure proper coverage of the interior surfaces. This early developmental work proved successful, and was the basis for applying clean-in-place operations to other pieces of equipment.

Generally, reference to clean-in-place applications infer that the equipment can actually be cleaned in place without any major disassembly of the equipment. With the development of larger pieces of processing equipment—particularly storage vessels—and the development of sanitary air-operated valves, sanitary products pumps, and welded product lines, the disadvantages of manual cleaning became evident.

Typical clean-in-place cycle. Basic clean-in-place operations in the food processing industry can be divided into one of three different cycles for line cleaning operations (which may include pumps, valves, or plate heat exchangers); vessel or spray cleaning operations (associated with tankers, tanks, processors, blenders, dryers, evaporators, and so forth); or the combination of the two in one program. Generally speaking, it is advantageous not to combine the cycles.

There are five basic steps associated with a typical clean-in-place operation: (1) pre-rinsing of the equipment with an ambient-temperature water to remove as much of the soil as possible, with the pre-rinse discharged to the drain; (2) recirculation of alkali or acid detergent solution at preset concentration and temperature; (3) post-rinsing with ambient temperature (or warm) water to remove wash solution residue, with the post-rinse discharged to drain; (4) rinsing of equipment with mild acidic post-rinse solution; (5) sanitization of equipment prior to use, not necessarily following clean-up of equipment.

This five-step sequence may be considered typical, but many variations are found in the field due to the type of clean-in-place recirculatory system utilized (that is, single-use, multiuse, single-tank, multitank, and so on); the type of equipment being cleaned; and the type of soil encountered. In some cases only an acid detergent may be required; in other cases, acid will be used followed by a caustic solution or such. Accordingly, for proper cleaning, consideration must be given to the use of appropriate chemicals, and the selection of a proper clean-in-place recirculating unit and associated controls, sprays, return pumps, and so forth, as part of the total system. In addition, the equipment to be cleaned should be designed for ease of clean-in-place applications. E. F. Reinke

Waste Management

Waste disposal is a matter of great economic and technical concern for most food processing operations, as well as other industries. Both solid and liquid wastes are encountered in the food processing industry. The disposal of these wastes is regulated by the Clean Water Act of 1977 (as amended) and related environmental protection legislation, which govern the types and amounts of waste discharged into the environment. As federal, state, and local environmental regulatory agencies increase their demands for less waste discharge, it becomes essential that cost-effective methods of waste disposal be found.

Waste management is a modern approach to cost-effective food processing waste disposal. Through waste management, modifications are applied to food plant operations and manufacturing processes. These modifications reduce the amount of solid and liquid wastes, recover more product and byproducts, often reduce energy consumption, and exhibit other benefits. In general, the principle is to convert waste liabilities into profitable assets.

One major objective of waste management is to eliminate or at least lessen the dependence on end-of-the-pipe sanitary engineering methods. This is achieved by reducing both the amount of waste solids generated and the volume of wastewater discharged.

Modifying plant operations. The following are examples of modifications which can be made to food plant operations: incorporating good housekeeping practices; collecting culls and other solid wastes into containers rather than discharging to the floor drain; recycling water; reusing spent process water in another plant operation; and using less or no water in plant operations that formerly used a fair to a large amount of water.

Good housekeeping practices that reduce water usage and wastes require good personnel management and employee awareness of conservation practices. Such practices as needless use of water or overloading of containers, thereby causing spillage, should be discouraged.

Recycling water in the same plant operation can be achieved by treating spent process water with activated charcoal or sand filters or by ion-exchange columns, chemical treatment, pH adjustment, temperature adjustment, pasteurization, or a combination of these and other methods. Brines are one of the more troublesome wastewater discharges to treat because of their high salt and organic content. Rather than discharging for wastewater treatment, spent olive, pickle, and cherry brines have been treated and recycled to start new batches of product (**Fig. 44**). Recycling of water has been used in other food plant operations.

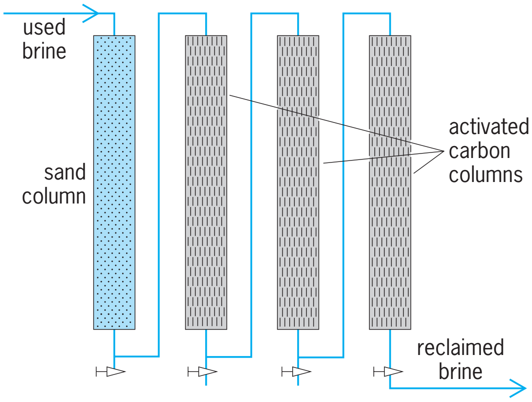


Fig. 44. Brine reclamation system used to recycle spent pickle, olive, and cherry brines for the purpose of initiating new batches of products. (*After J. H. Green and A. Kramer*, *Food Processing Waste Management*, *AVI Publishing Co.*, *1979*)

Countercurrent water reuse systems can be established in many plant operations. For example, potable water used to rinse the finished product prior to packaging can be used a second or third time to wash the product during intermediate stages of processing and to wash the incoming commodity. The spent wash water could be used again to initiate washdown of dirty floors or to flume solid wastes away from the processing line.

Examples of methods that use less water or no water in plant operations compared with former methods are dry fluming such as mechanical or pneumatic rather than hydraulic transport systems; dry cleaning using shakers or brushes to initially clean dirt from commodities or equipment; and using high-pressure, low-volume cleaning wands for plant and equipment cleanup. The last example also reduces the amount of time, labor, hot water, and detergents needed for plant sanitation operations.

Process modifications. These often achieve the greatest reductions in water usage and solid waste discharge; they may also exhibit other benefits such as energy conservation. Process modifications require a large amount of time, technical input, and money to do the necessary research and development. Federal agencies such as the Environmental Protection Agency (EPA) and the U.S. Department of Agriculture (USDA) have funded projects through their own researchers or through universities and trade associations. Trade associations such as the National Food Processors Association (NFPA), in cooperation with industry and federal agencies, have been leaders in the research and development of food processing modifications. The return on investment has been outstanding for many of these process modification projects.

Two of the best examples of food processing modifications were the development of dry (mechanical) peeling methods (late 1960s) and steam peeling methods (1970s) for potatoes, other root crops, tomatoes, and tree fruits. For many years (prior to the late 1960s) potatoes and other root crops were peeled by immersing them in hot caustic solutions and removing the treated peel with blasts of jet-sprayed water. In the dry peeling processes, hot caustic or steam-treated peel is mechanically removed by rubber or neoprene brushes and rotating disks. In the late 1970s, steam peeling methods were developed for root crops and tomatoes. The advantages of dry peeling are that no water is used in peel removal; the peel waste can be collected separately for by-product development; often more product (less waste) results; less volume of hot caustic peeling solutions is required; and the pollution waste loads are greatly reduced. This process modification has been shown to be very cost-effective, and it conserves energy.

Modifications of the blanching process have reduced the tremendous waste loads that often resulted from the use of large volumes of hot water in blanching tanks. Modifications to blanching include individual quick blanching, steam blanching, or hot-air blanching. Mechanical innovations such as the vibratory spiral blancher-cooler developed by USDA have increased the efficiency of steam blanching. These modifications have resulted in reducing waste loads, retaining more nutrients in the product, and conserving energy.

Other examples of process modifications include the following: deboning to recover more product from meat, poultry, and fish; ultrafiltration followed by reverse osmosis to separate cheese whey into proteins, lactose, and minerals; ultrafiltration or reverse osmosis or both used in other applications to separate components for by-product development; use of heat exchangers and other process modifications in the citrus industry to save energy, reduce pollutant loads, and recover by-products. The fishing industry has its own particular waste disposal problems. The traditional methods of fishery waste disposal have been throwing the waste overboard from fishing vessels and tossing the waste over the end of the wharf by the processing plant. These traditional methods of fishery waste disposal have probably fostered a lack of effort by this industry to find ways of utilizing fishery wastes or to find methods of improving processing efficiency. Local and national concerns have discouraged these traditional methods of fishery waste disposal to protect the environment. Solid waste losses from processing fish and shellfish have ranged from about 30 to 85% depending upon the type of fishery: fin fish filleting operations generate 30–60% losses, shrimp processing operations generate 40–80%, and crab processing operations generate 75–85%. There is a recognized need to improve resource recovery from fishery processing operations, and this has generated some specialized waste management operations for the industry.

Mechanical deboning of underutilized species, or fish processing wastes, has been carried on for many years in some segments of the industry. In the late 1970s, interest in the Japanese methods for using surimi (deboned fish flesh) was transferred to the development of a wide variety of new products from deboned fish. Thus former wasted fish flesh from underutilized species of fish, or on the frames of filleted fish, can now be converted to many new products.

The traditional fish meal industry is capital-, energy-, and technical-labor-intensive. It has been replaced in some parts of the world by the fish silage (liquefied fish) process. Industrial fish, underutilized fish, and fishery wastes can be treated with organic or inorganic acids, or combinations of both, to form fish silage, an animal feed product. It is formed by the autolytic enzyme breakdown of fish proteins. The fish silage process does not require large amounts of capital, energy, or labor and is therefore well suited for small as well as larger fish processing operations for the purpose of developing by-products from wastes. However, fish silage is about 80–85% water, and costs of transportation becomes a limiting factor for its use. If there are no local animal feed markets for fish silage, it can be combined with cereals, dried together, and shipped to distant markets.

By-product development. Recovered food-processing waste solids offer the opportunity for development of food, feed, fodder, fiber, and fuel by-products. In addition to the examples of by-product development given above, there are other innovative methods such as microbial technology used to develop single-cell proteins, fuels (for example, alcohol and methane), enzymes, and other organic chemical products from food-processing wastes and wastewaters.

Some solid wastes such as olive pits, nut shells, and bagasse (sugarcane waste) are being used as fuels. Many sugar mills in Hawaii and Brazil burn bagasse to generate steam and electric power to operate the sugar mill. Often excess electric power is generated, which is sold to the electric utility company. *See* COGENERATION SYSTEMS.

Shellfish wastes contain a large amount of chitin in addition to protein and calcium carbonate. Chitin, which constitutes much of the exoskeleton and some of the internal parts of shellfish, is essentially inert to most chemical reactions and therefore has little industrial use. However, shellfish wastes can now be transformed into useful products (**Fig. 45**). The wastes are treated with dilute sodium hydroxide and dilute hydrocholoric acid to produce proteins, calcium chloride, and chitin. Then a hot concentrated (40–50%) sodium hydroxide treatment deacetylizes the chitin, forming chitosan, which has a wide variety of uses and is therefore a valuable product. Shellfish processing wastes are transformed into chitosan and other by-products (protein and calcium chloride) in Japan and some other countries.

Sanitary engineering. Although the objective of waste management has been to greatly reduce or eliminate the need for end-of-pipe treatment, sanitary engineering methods are still applied to many food processing waste effluents. Both primary sedimentation and secondary biological sanitary engineering treatment methods have been applied to food processing waste effluents. Typical secondary treatment methods are: various forms of activated sludge treatment, including the oxidation ditch method; trickling filters, including the rapid and super rate modifications; anaerobic, aerobic, and aerated lagoons; and land application of wastewaters such as crop irrigation, infiltration-percolation, and overland flow methods. The activated biofiltration

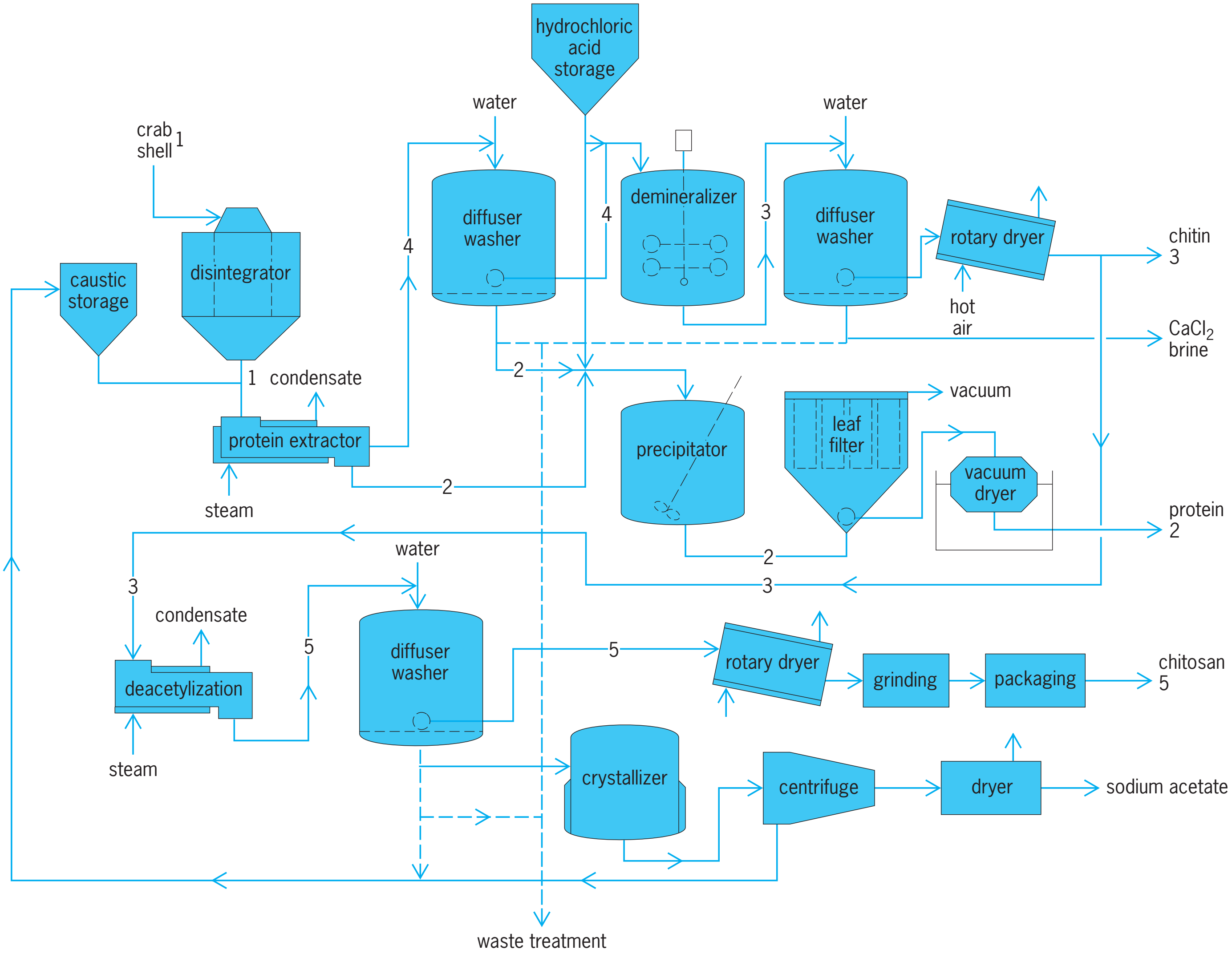


Fig. 45. Chitin-chitosan process flow chart. Process number (1) indicates waste crustacean shell, (2) protein, (3) chitin, (4) deproteinized crustacean shell, and (5) chitosan. (*From J. H. Green and A. Kramer, Food Processing Waste Management, AVI Publishing, 1979*)

method, a combination of trickling filter and activated sludge technologies, has proved to be very effective in treating fruit and vegetable wastewaters.

Many end-of-the-pipe sanitary engineering methods are both costly, in terms of capital investment and operational costs, and nonproductive. Processors that discharge into available municipal sewer systems will encounter increased service charges and probable requirements for pretreatment prior to discharge. On-site sanitary engineering installations can be modified to recover solid wastes for by-product development. Some food processors are processing sludge recovered from activated sludge treatment of food wastes into animal feed. Treated wastewater can be used for crop irrigation or for recycling as process water.

Food processors, as well as other manufacturers, had to consider tertiary or advanced wastewater treatment methods in order to meet the EPA's 1985 zero discharge effluent requirements. As an alternative to advanced wastewater treatment, many food processors are considering land application of primary or secondary treated wastewaters. Food processing wastewaters, which do not contain toxic substances, can be used to grow crops or recharge ground water. There are three methods of land application used for food processing wastewater (**Fig. 46**): application to crop land (clay to sandy, loam) to irrigate crops or orchards; applications to sandy loam or sandy soil to recharge groundwater; and overland flow to clay or clay loam soils planted with grass ground cover. In the last, the grass sets as a filter system and holds back waste solids. The harvested crops provide monetary return on the costs of waste disposal.

A food technology profession has emerged, known as waste or resource recovery management. The waste manager is closely allied with (if not the same person as) the energy conservation manager. The challenge is to find ways of reducing water usage, recovering waste solids, developing marketable by-products, conserving energy, and

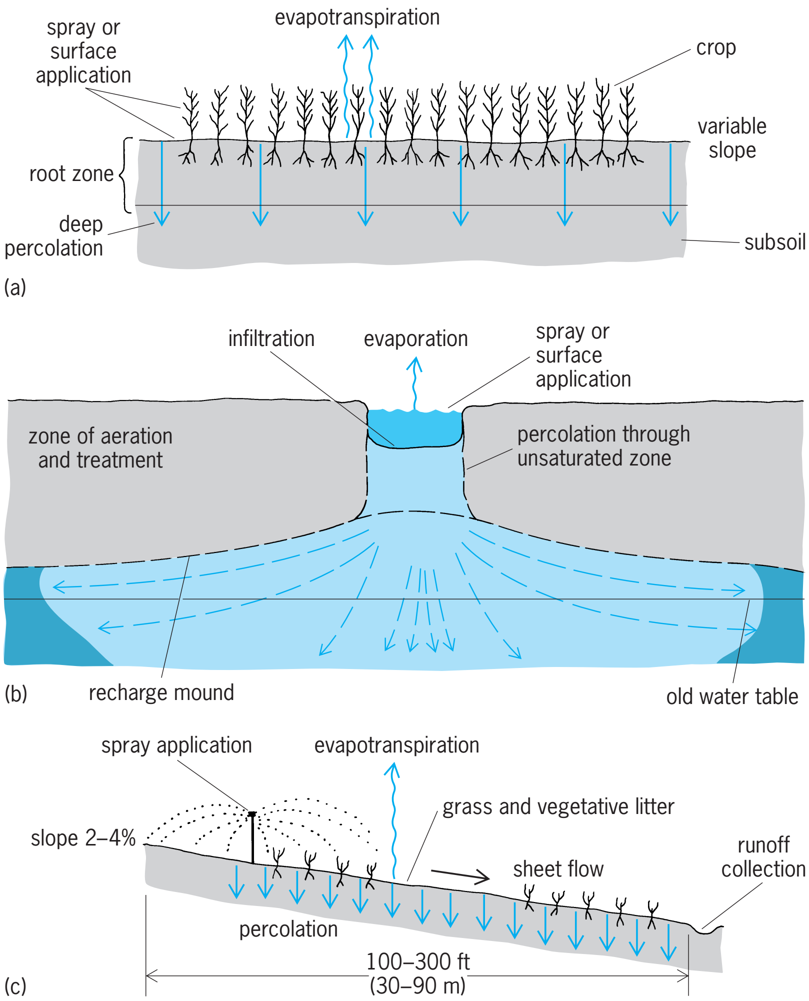


Fig. 46. Land application of pretreated food-processing wastewater. (*a*) Irrigation; sandy-loam soils to raise crops. (*b*) Infiltration-percolation; sandy soil to rechange ground water. (*c*) Overland flow; graded clay soils planted with grass or other ground cover. (*After J. H. Green and A. Kramer*, *Food Processing Waste Management*, *AVI Publishing Co.*, *1979*)

finding alternatives to costly, nonproductive end-of-the pipe treatment methods. *See* ENVIRONMENTAL ENGINEERING; SEWAGE TREATMENT; WATER POLLUTION.

John H. Green

Process Control

In what was probably the first food engineering application of feedback techniques, Oliver Evans in 1784 used control techniques to automate a Philadelphia flour mill. Correctly applied, control instrumentation can give a product closer to specifications than manual control with reduced waste and increased yields. Advantages of automated process control include higher efficiencies with maximum utilization of all ingredients and raw products. *See* AUTOMATION.

Control tasks. The implementation of process control is based on the particular control task. There are four fundamental control tasks in industrial food processing: sequential control, loop control, interconnected control, and adaptive control. *See* CONTROL SYSTEMS.

Sequential control. This is for the independent batching or sequence operations. Most processing of food is derived from familiar domestic patterns that are performed in sequential batch-type steps. Controlling elements in such sequential control paths are typically handled by programmable logic controllers. *See* PROGRAMMABLE CONTROLLERS.

Loop control. This control task is based on a feedback or feedforward information system. The control loop is fundamentally an algebraic summing point between a set point (desired value) and the measurement (current value). The corresponding error is used as a means of implementing a specific corrective response. The most common type of control loop is a feedback control loop. Basically the feedback controller reacts to the output from the process and regulates the process based on that output condition. Feedforward control, an alternative approach, refers to the control action in which information concerning all the conditions that can disturb the variable being controlled is converted into a specific corrective action in order to minimize deviations of the output.

Of the four general control tasks, loop control is the one most utilized within the food industry. It is a fundamental component of the interconnected and adaptive control tasks, and is usually found as an associated control system within sequential control tasks. There are four critical elements to any loop control system: (1) the process to be controlled, (2) the measurement of the current conditions using sensors, (3) the controller, (4) the final output from the controller affecting the system.

All of the techniques grouped together under the term loop control have one feature in common, and that is a continuous path from sensor through a controller and then to an output. This path can be broken intentionally, thus producing an open-loop condition for study of the controller or maintenance of the system.

Interconnected control. Under this type of control, the controller has to adapt its output, because when it controls one variable it also, by the action of changing or altering that variable, is altering a second controlled variable. A condition called coupling exists when a change meant to alter one variable also alters another variable. Interaction exists if there is coupling the other way around.

Adaptive control. A control task typical of applications in food processing is sometimes called adaptive control. Under adaptive control a model of the system is constantly tested and updated based on raw product input and processing conditions. Changes are implemented automatically in the control strategy by adjusting various parameters within the process system. There are many opportunities for this type of modeling and optimization within the food processing industry. *See* OPTIMIZATION.

Single-loop controller modes. There are basically four types of common control modes within single-loop controllers. How a decision is made concerning the implementation of a specific control mode depends on a number of factors, including economics, precision of control required, time response of the process, processing characteristics, and safety considerations.

Two-position control. Within this classification there are three common control options. On-off control produces a control function based on a measured variable, comparing it to a set point. When the measured variable crosses the set point, the output is altered to the opposite position. The output can be in only two positions: full on or full off. The second type of two-position control is a differential gap controller, whose output is based on the time a measured variable crosses an upper and lower boundary. The third type is a time cycle system. Typically in such a controller a specific time cycle is established in which the final output must complete a full cycle; that is, if it starts in an open position, it must close and then reopen during that cycle. The time in which a valve occupies the position opposite to that it occupied at the start of a particular time cycle depends on the measurement point and whether it is above or below the set point.

Throttling control. The three other control modes are based on proportional or throttling control. Throttling control indicates any type of control system in which the final output is purposely positioned to achieve a balance between the supply and demand of the process. This basic type of throttling control is called proportional-only control. Under this type of control the position of the final output is based on the error between a set point and the measured variable multiplied by a gain factor and added to a bias. Proportional-only control attempts to return a measurement to a set point after a load upset has occurred. However, based on its design, it is impossible for a proportional controller to return the measurement exactly to the set point.

The difference between the measurement and the set point after a proportional-only controller has come to equilibrium under new load conditions is known as offset. If the offset cannot be tolerated, an additional control function must be added to the proportional-only control resulting in the third type of controller, proportional-plus-integral control. Integral action will integrate the difference between measurement and set point over time and cause the controller's output to change until that difference is brought to zero. The basic limitation of this proportional-plus-integral controller is called integral windup. This must be seriously considered on discontinuous or batch processes, because under these processes the controller output can be saturated and overshoot the set point. A mechanism must be integrated into the controller system in order to prevent this overshoot.

The fourth and final type of controller takes into account throttling control, maintenance of zero offset, and control action based on the rate of error change. This controller is known as a proportional-plus-integral-plus-derivative controller. Derivative action occurs in addition to the proportional and integral action whenever the measurement signal is changing. The derivative action is proportional to the rate of change in the measurement signal. By reacting to the rate of input change, the derivative portion of the controller allows the unit to inject more corrective action than is initially necessary in order to overcome system inertia.

Sensors. In any process control situation, accurate and reliable measurement of process conditions are critical in order to implement an adequate control strategy. A lack of suitable sensors for many applications limits the progress of automation in a number of areas. Producing sensors of sufficient accuracy, reliability, and ruggedness for operation under plant conditions is many times more difficult than designing equipment for making the same measurements under laboratory conditions. Measurements commonly made in the food industry include temperature, flow, and level. Some more specialized requirements involve measurements of viscosity, pH, conductivity, humidity, density, and color.

Process regulators. Every control system has a means for switching or regulating process equipment by some sort of final control element. There are a number of devices available for convenient on-off switching. Power switching employs relays, contactors, and solid-state relays; pneumatic and hydraulic power use solenoids. Reliable and efficient regulation of speed or flow on a continuous basis is difficult to accomplish simply with on-off devices.

Continuous variation of temperature, speed, flow rate, and other process conditions is generally achieved with one of four types of regulators: flow-controlling valve, variable-speed motors, mechanically variable pumps, and variable-power electric heating. The flow-controlling valve is the most commonly used regulator. It usually requires special sanitation precautions, and it is generally inefficient because it imposes pressure drops in the control lines. However, despite these disadvantages, it is fairly popular because in its manual form it is very simple and direct. Generally it needs only one moving part.

Pneumatic linear accuators are popular for driving valve stems. Simple types use air on one side of the diaphragm working against a spring; other types use pneumatic position feed. Pneumatic signaling with a working range of 3–15 lb/in.^{2} (21–103 kPa) has been widely used for many years in virtually all pneumatic control systems. Where computers and electronic integrated circuits are installed for performing the control function, pneumatic accuators may still be used through the use of an electronic transducer. This transducer converts digital or analog values to specific pneumatic levels within the working range. *See* SERVOMECHANISM; TRANSDUCER.

Direct digital control. A significant advancement in the microminiaturization of digital electronics accompanied by a reduction in cost has facilitated the use of digital logic in all control tasks. Computerization of process control strategies allows development to progress from sequential startup through transition into continuous operation on a virtually 100% automated basis.

The availability of systems with inherently large data processing capabilities may lead to process automation in small plants or automation by stages, as well as total plant control. Very sophisticated small computers will permit implementation of a very diversified, distributed control strategy, with flexibility and versatility. *See* COMPUTER-INTEGRATED MANUFACTURING; DATABASE MANAGEMENT SYSTEM; DIGITAL COMPUTER.

Mark R. McLellan

Packaging Equipment

Packaging methods are influenced by the type of product and the packaging concept, and include filling, closing, labeling, multipackaging or unitizing, and coding. Primary packaging involves a close partnership between the product to be contained and the packaging material in contact with the food or beverage. Secondary and tertiary packaging encompasses collation or unitization of a number of different primary or secondary packages.

Primary packaging equipment. Packaging lines for preformed packages usually contain components such as accumulators, unscramblers, infeeds, aligners, fillers, closers, and labelers for preformed containers which include cans, glass, bottles, tubs, trays, and similar items. Prior to the infeed, the preformed containers are received from the fabricator, or from an in-plant forming operation, and introduced into a line. Before filling, the containers are cleaned to remove foreign matter resulting from fabrication and transport and to ensure their general cleanliness.

Metal cans are generally received in bulk shipment aligned in tiers on a pallet; glass jars may be received in corrugated fiberboard cases or in palletized or bulk form; and plastic jars and bottles may be received in cases, on pallets, or in dump bins. Preformed plastic trays, cups, or tubs are usually received nested in stacks. Preformed bottles or cans that are received in palletized form are removed from the pallets by mechanical depalletizers such as sweepers or vacuum cup pickers which move the containers onto accumulator tables that feed aligners. Plastic bottles or jars that are received in bulk are unscrambled, righted, and conveyed from an accumulator table to the infeed of the packaging line. Foreign matter within the containers may be removed by air-pressure cleaning or by inverting the cans or jars and subjecting them to air blasts or jets of water followed by air. Cans, bottles, or jars received in cases are mechanically removed from the cases. Trays, cups, or tubs are mechanically denested.

Paperboard folding cartons are received as precut flat or preglued blanks which must be formed or erected. Flat blanks may be erected by applying adhesive to the corners and then forcing the blanks into a mandrel to bring together the walls and tabs or edges to form an open-top or open-ended container. Another common method employs die-cut corner-locking tabs; this requires mechanical locking to erect the trays. Many paperboard cartons are received as preglued tubes which are erected by vacuum removal from a magazine and snapping open. Using preglued sleeve-style blanks, an automatic carton machine (**Fig. 47**) erects the carton, conveys the open carton for product insertion, and seals the carton ends by activating a heat-sealable coating on the paperboard surface. Typical speeds are 70–300 per minute.

In the type of double-package maker that is widely used for dump filling of dry cereals, cookies, crackers, and snack foods, an internal flexible coextruded polyolefin liner is formed around a mandrel after which a paperboard flat blank or a tube is wrapped around the liner on the mandrel. The carton with its inner liner is glued into position, and the entire preformed package is stripped from the mandrel or to a line in which it is filled and sealed.

Flexible pouches are usually formed in line with the filling equipment. Form/fill/seal equipment includes both vertical and horizontal types.

In vertical equipment, a web of flexible material is unwound and turned over a forming collar into a vertical tubular form where the edges are brought together and heat-sealed into a longitudinal tube. While the tube is in the vertical attitude, a transverse seal is formed at a base, the product is filled by gravity feed which may be assisted at the outset, and a transverse seal is made at the top. In most such units, the transverse seal for the top of one pouch constitutes the seal for the bottom of the next pouch in line. Vertical machines generally operate at cycle speeds in the range 30–40 per minute for packages of about 2–6 oz (60–170 g) contents. Output speeds may be increased by multiplying the number of tubes in parallel with each other. Output speed and reliability have been increased by incorporating drive belts on the side of the web on the tube to assist the downward motion of the tub. Vertical form/fill/seal machines are used extensively for snack foods, dry granular food products such as nuts and candies, and free-flowing products such as coffee. In multiple-lane form with face-to-face heat sealing, this system is used for very high-output unit-portion packaging of fluids such as mustard, ketchup, or mayonnaise and hot filling of high acid fluids such as tomato sauce. With chemical web sterilization, vertical form/fill/seal constitutes the principle for aseptic packaging of liquid food products such as milk and juices in brick- or block-shaped paperboard–aluminum foil–polyolefin laminations.

Horizontal form/fill/seal packaging equipment also operates from roll-stock flexible packaging material. With continuous motion, multiple filling heads in parallel with the forming, speeds of up to 700 packages per minute may be achieved, particularly for smaller-type pouches such as unit-portion sugar packets. The horizontal packaging systems are employed to package dehydrated soups, powdered drink mixes, and dry condiment and seasoning mixes, but speeds in these larger sizes are only in the range of 60–120 packages per minute, generally in intermittent motion mode.

Stand-up pouches with an added section of flexible web on the base, and gusseted sides, have been available since the 1960s, but only in the 1990s did they gain importance. Environmentalist pressures have suggested stand-up pouches as substitutes for higher-mass semirigid plastic bottles for liquids or for folding

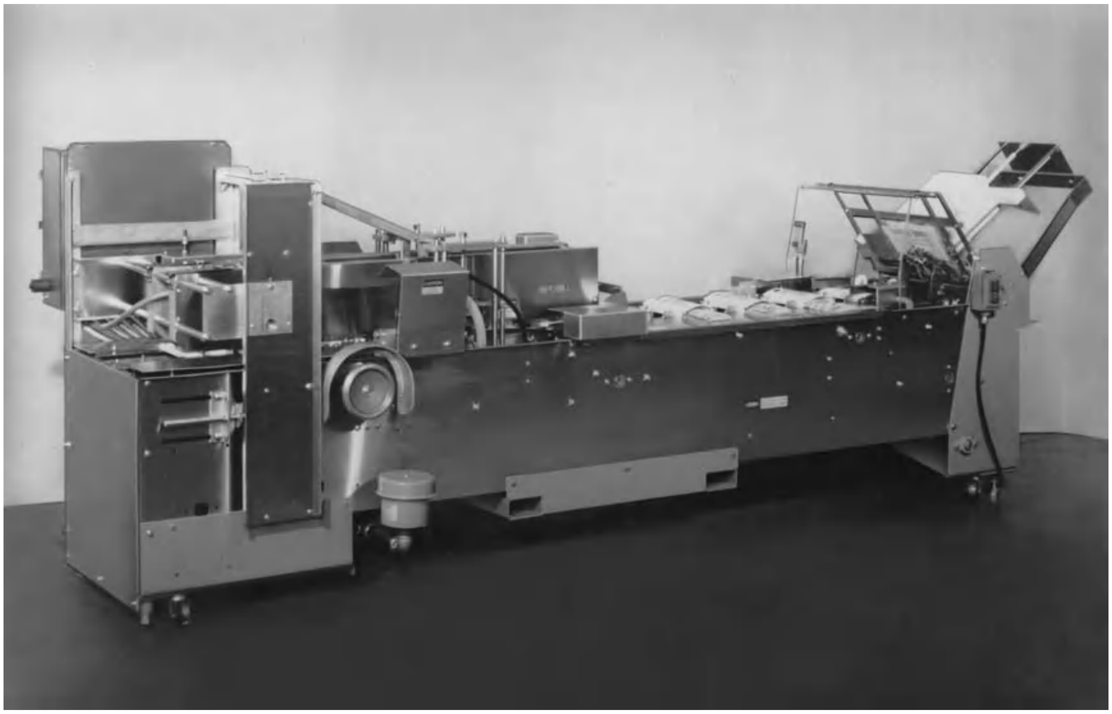


Fig. 47. End-load-style cartoning equipment using preglued sleeve style blanks. (*Kliklok Corp.*)

cartons for dry foods. Most stand-up pouches are prefabricated, with a separate web used to form the bottom gusset. Some, however, are formed on modified horizontal form/fill/seal equipment by incorporating a W-shaped fold in the base to effect the stand-up feature. Although they are accepted and used in Japan for liquid foods and in Europe for liquid household chemicals, their application in the United States has been confined to dry products such as fruit and meal-item mixes.

Fillers. Filling components are designed and engineered for specific categories (dry granular, dry solid, liquid, and viscous fluid) of food products. No single filler design is capable of handling the full range of food products. Dry-food-product fillers may have speeds of over 1000 packages per minute, provided the packages are very small size. Typical filling speeds for liquids for cans and bottles range from 60 per minute for low-speed juice fillers, through 600–900 per minute for individual unit-portion cans or bottles of juices, up to 1500 per minute for beer or soft drinks with equipment now operating at up to 2000 per minute for beer cans. Filling speeds may be increased by multiplying the number of filling heads, counting stations, weighing stations, and so forth.

Basic types of filling machines for dry products include net weighers that weigh the product prior to discharging into the package. Electronic scales operate with significantly greater accuracy and precision than older mechanical equipment (**Fig. 48**). Computerized fillers select from a multiplicity of small preweighed shots, and calculate the specific individual shots to be summed to the whole that becomes the package fill.

Gross weighing devices scale the product into the package, and employ the gross weight of the product and package to determine the quantity to be filled. Volumetric fillers are used for both liquid and dry products. The product, which has a predetermined density, is placed into a fixed volume container above the open-top package. When the receptacle has been filled, the contents are discharged into the package. Metering fillers establish a standard rate of product

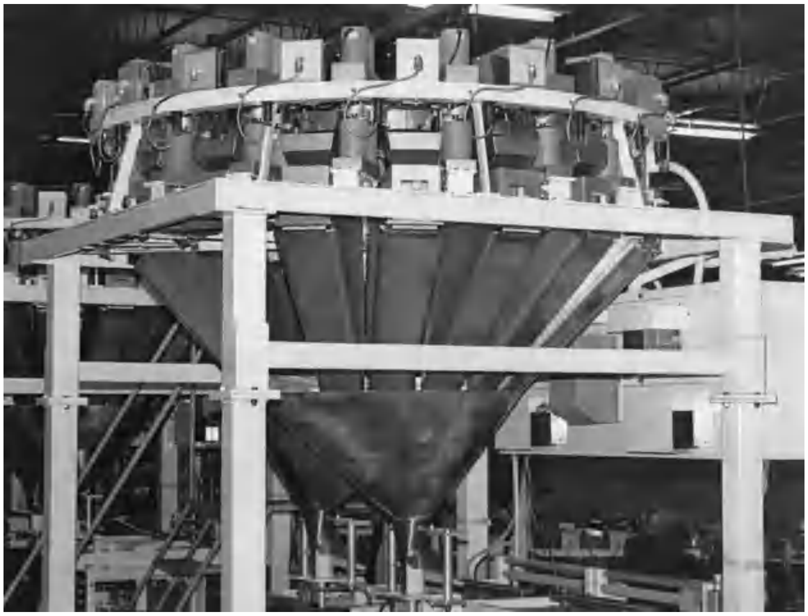


Fig. 48. Computer-controlled multiple-scale weighing device for snack foods used on vertical form/fill/seal flexible packaging machine. (*Kliklok Co.*)

flow for filling into the preformed package at specific time intervals. Auger-type fillers rotate a specified number of times to force the powdered or granular product, such as instant coffee or tea, down through an orifice and deliver by gravity the measured amount of product. Other fillers may count the number of units (such as tablets) being placed in the package.

Bottles are filled with liquid by gravity, gravity vacuum, or pressure gravity. In gravity filling, liquid flows directly by gravity from an overhead surge tank through the filling valves with the nozzle cut off at a predetermined level within the bottle. In gravity vacuum, the entire bottle neck is in a vacuum system, and the force of gravity is supplemented by mechanical vacuum drawn on the bottle interior during filling. Pressure gravity systems employ pumps to increase liquid pressure, especially for filling more viscous liquids. In vacuum filling, air is evacuated from the bottle to assist gravity, and atmospheric pressure above forces liquid from a reservoir.

Filler types for viscous fluid products are either straight-line or rotary-plunger. In either case, cylinders and reciprocating pistons are employed to accurately measure and force the product under pressure into the containers below. Positive-displacement fillers are employed for viscous products such as cottage cheese, ice cream, and juice concentrate; for filling bottles, cans, or paperboard containers with liquids such as milk and juice; and for aseptic filling of liquids in horizontal machines. A vertical form/fill/seal liquid packaging system involves a tube of preformed paperboard lamination packaging material wrapped around a liquid already present. The package is sealed through the liquid to deliver a pouch that is converted in-line into a brick shape.

Closing and sealing. Bags or pouches incorporating interior thermoplastic materials are usually sealed by the application of heat and pressure. Heat sealers include roller, band, bar, high-frequency, and impulse sealing. Induction heating requires the presence of metal in the structure to achieve sealing temperature. Most flexible materials and paperboard-foil-polyolefin laminations are closed mechanically and heat-sealed by fusing the polyolefin coatings to each other under pressure.

Paperboard cartons are closed by mechanical locking, wet glue, hot-melt adhesive, or occasionally by fusing the interior coating to itself. Metal cans are sealed almost exclusively by rotary double seamers that apply rigid metal tops, and mechanically clinch the top closure to the can body over a flange.

Can, glass bottle, and jar closers and sealers are usually off-line from the fillers and may be rotary or straight-line. Automatic capping and closing machines are employed to unscramble or align lug-type caps, closures, side-seam closures, and continuous-thread caps.

Labeling. Food cans are usually labeled in the horizontal attitude by rolling them over stacks of paper labels which have had adhesive applied. In one special method, roll-stock paper labels have been applied to cans at very high speed with die cutting between each can.

Many beverage and food bottles are prelabeled by using preprinted foamed polystyrene applied in the glass plant or by imprinting the bottles with color. Other bottles and jars require paper or other composition, for example, polypropylene labels which may be full wraparound or partial labels or partial-plus-neck labels. Continuous-motion labeling can be matched to the speed of a bottling line, that is, well in excess of 1000 per minute. Pre-die-cut labels are removed from a stack, adhesive is applied, and the label is wiped on to the bottle which is passing by at high speed with axial rotation.

Pressure-sensitive labels with preapplied adhesive are applied either mechanically or by air blowing from roll-fed stock onto the packages.

Bag-in-box. Many food products are packed in bulk bag-in-box for industrial and commercial applications. Bag-in-box had been used for many years for commercial chilled milk packaging, and this led to development of bag-in-box packaging for dry products such as nuts, fruit, frozen fruits and candy, aseptic packaging for juices and juice purees, packaging of wine, and hot filling. The bag is usually prefabricated and inserted into a corrugated fiberboard case with a pre-die-cut opening for the valve through which the product may be filled. Hot-fill pouches are generally prefabricated, filled through an open top, and closed by heat sealing.

For products requiring further protection, such as fruit purees, the bag is fabricated from a barrier material. Metallized polyester or aluminum foil are now employed to effect the oxygen barriers demanded by the contents.

Single-wind paperboard cans. The single-wind convolute paperboard can is designed for in-plant package fabrication to produce cylindrical containers capable of containing dry granular food products. The systems begin with preprinted flat blanks, thus eliminating the need to introduce preformed containers and saving on space volume. Equipment provides the means to form the containers in-line with the filling operations.

Overwrapping systems. Wrapping machines for flexible materials may be employed directly on the product or may be used to wrap other primary packaging for what is perceived by the packager or consumer to be added protection. Direct wrapping in flexible materials is common for products such as soft baked goods, fresh meats, and cheeses.

Wrapping machines may employ either reciprocating or continuous motion. In one variation, the product may be inserted into a flight section which is then elevated to the level of the precut flexible material. The material is sealed using the product as a mandrel. Typical speeds for wrapping machines vary from 60 per minute for double-point end foldup to up to 300–600 for horizontal fin seal form/fill/seal wrapping that is used for unit-portion cookies and crackers. Most overwrapping is horizontal form/fill/seal.

Thermoform/fill/seal operations employing two webs—one a semirigid or flexible plastic material and the other flexible—are widely used for hermetically packaging processed meat, cheese, jams, jellies, and liquid dairy products. In horizontal thermoform/fill/seal packaging systems, the semirigid or flexible base plastic web is first heated to soften the material and then pressure-formed into a cavity. The web containing the cavity is moved to a filling station where the product is inserted. The second flexible web is brought down over the filled cavity, and the web is heat-sealed to flanges surrounding the cavity. Cavities for processed meats and cheeses may be evacuated by vacuum or gas-flushed to extend the refrigerated shelf lives of these products.

Secondary packaging equipment. This type of equipment is designed to accumulate a number of primary packages to ensure that they are protected in the distribution cycle. Secondary packaging equipment permits handling a multiplicity of primary packages as a single unit, and thus increases the efficiency of ditributing packages. It also permits the consumer to purchase more than one primary package at a time. Usually, secondary packaging is a module integrated downstream from primary packaging.

Multiple packaging of bottles may range from the relatively simple drop-into-paperboard basket carriers to the type in which the semirigid plastic or glass bottles are collated and placed into a unit that wraps a paperboard blank around the bottles and locks or glues it. Multiple packaging of bottles may also be accomplished by collating the bottles into groupings of four, six, or eight and bringing an all-plastic clip over the top to unitize the packages.

Cans may be multipacked by aligning them in rows of two or three and bringing a series of polyolefin ring carriers over the top and parallel with the movement of the cans. At appropriate intervals, the plastic ring carriers are stretched over the top of the chimes (end seams) of the cans to capture a grouping of four, six, or eight cans and thus form a multiple.

Paperboard cartons are employed to multipack 12, 18, 20, 24, and so forth cans or bottles. The most widely employed 12/24 multiple packers use preformed paperboard tubes. The tube is removed from the magazine stack and erected in a continuous motion. Groupings of 6 or 12 are pushed in from each side to form 12 or 24, respectively. Alternatively, a flat paperboard blank may be wrapped around a preformed grouping of 12, 24, and so forth cans and then either locked or glued into place. One system groups 12 bottles and wraps and glues a flat paperboard blank around the grouping.

Printed shrink-film wrapping has increased in importance for multipacking. In continuous motion, cans are grouped into 6, 12, and so forth counts, and a sheet of printed film is cut from a web, conveyed to the grouping, and wrapped over the top. A hot-air tunnel shrinks the film snugly to the unitized cans.

Tertiary packaging equipment. Most tertiary packaging in the United States is in corrugated fiberboard cases or trays. Loading of primary or secondary packages may be by dropping into the top or side loading of preformed units, or wrapping flat blanks around a collated block of primary or secondary packages, all followed by glue closing. Shrink-film bundling used imprinted shrink film to unitize trays of 24, and so forth, cans, usually on continuous motion equipment. Reciprocating-motion shrink-film bundling machines operate at speeds of up to 30 cases per minute. There are shrink-packaging machines which operate at speeds of up to 90 cases per minute to collate and wrap groupings of cans or bottles in polyethylene shrink film.

Stretch-film palletization of cases has become very significant. After tertiary packaging, cases are generally palletized for unit distribution. The film binds the cases tightly without further damage and permits relatively safe distribution of full pallet loads of food products. *See* FOOD; FOOD MANUFACTURING.

Aaron L. Brody

Aseptic Packaging

Aseptic processing and packaging is employed in the food and drug industries to package pathogen-free products that will not spoil when stored on a dry shelf. Aseptic processing is the treatment used on the product to eliminate pathogens and other microorganisms capable of causing product spoilage during ambient storage. A packaged product that does not contain viable pathogens or other

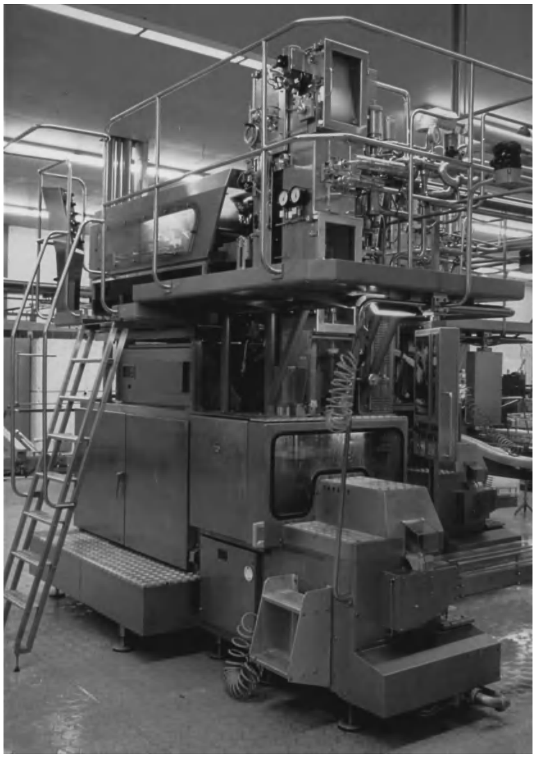


Fig. 49. Aseptic packaging machine for forming brick-shaped packages. (*Tetra Pak, Inc.*)

microorganisms capable of product spoilage is considered commercially sterile. Aseptic packaging is the process of filling sterile product into sterile containers under conditions where viable microorganisms are prevented from recontaminating the product. Commercial aseptic packaging machines are sophisticated systems capable of sterilizing the container material, forming the container, filling the product with consistency in fill weight or volume, maintaining sterile conditions while filling, and sealing the package securely before it leaves the sterile section of the machine. These operations must be carried out at high speed for the process to be commercially viable.

Machinery. Aseptic packaging machines vary in operation and type of package produced. Some receive containers that have been sterilized at another location. These machines must permit transfer of the sterile containers into the filling section without recontamination. The inability to detect microbial incursion at any time during the transfer and filling process is a serious limitation. Thus, when packaging low-acid food products, which support the growth of the deadly microorganism *Clostridium botulinum*, packaging material sterilization and package forming, filling, and sealing, are integral parts of the same machine. Most aseptic packaging machines are preprogrammed to go through a presterilization process before packaging operations can begin, and operating conditions are monitored by using appropriate sensors for temperature, pressure in the filling section, package sterilant temperature, and flow. An automatic shutoff system, which would terminate filling when conditions indicate probable loss of commercial sterility, is an important component of aseptic packaging machines.

Sterilization. When the packaging material is heat resistant, such as metal and glass, the packages are sterilized in-line with superheated steam at 300°F (150°C) or higher. Glass must be heat sterilized with care by prolonging time to reach sterilization temperature and by cooling down slowly before filling to prevent breakage. Plastic packaging materials are sterilized by dipping in, or by spraying into the formed package, 15–35% hydrogen peroxide solution. Hydrogen peroxide is removed from the sterilized package by drying with radiant heating elements and a stream of hot air. Regulations by the U.S. Food and Drug Administration (FDA) limit the level of residual hydrogen peroxide in the product to 0.1 part per million. Hydrogen peroxide removed from the packaging materials after sterilization can create very uncomfortable conditions near aseptic packaging machinery. Machines must prevent leakage of the air–hydrogen peroxide mixtures into the surroundings. Hydrogen peroxide is removed from air by condensation before the air is vented out through a stack. Other sterilants, such as ethanol and high-intensity ultraviolet light, have been suggested and are used to a limited extent in other countries. However, in the United States hydrogen peroxide is the only chemical sterilant approved by the FDA for

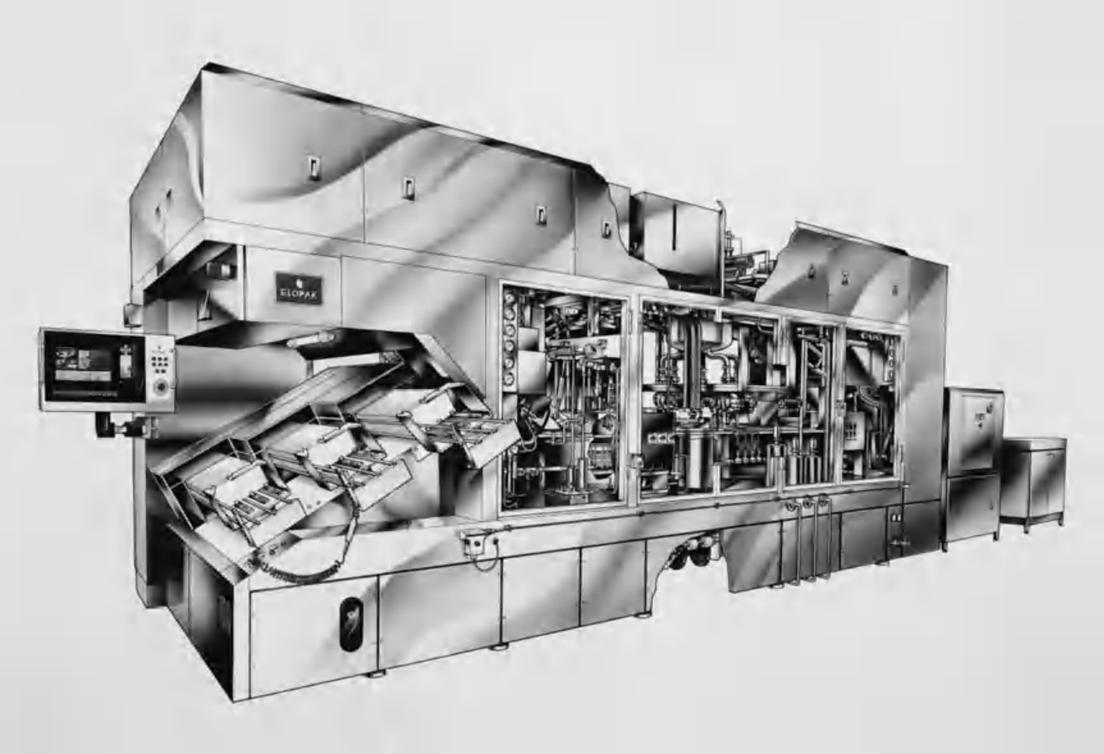


Fig. 50. Aseptic packaging machine for forming gable-top containers. (*Pure-Pak, Inc.*)

use in aseptic packaging machines. Packaging materials may be presterilized in a separate location with gamma irradiation, overwrapped to maintain sterility during transportation and storage, and transferred out of the overwrap by using an elaborate device to prevent recontamination.

Filling. Filling must be done in a presterilized filling station. Machine sterilization is accomplished by spraying hydrogen peroxide solution inside this section of the machine for a specified time, followed by drying with hot sterile air. Some machines are presterilized with superheated steam or air at very high temperature. Sterile air continuously admitted into the machine to maintain a positive pressure maintains sterile conditions during the operation. Filters capable of screening out microorganisms are used to produce the sterile air.

In the pharmaceutical industry, sterilization of clear injectable solutions is accomplished by filtration. Aseptic packaging is accomplished by using presterilized containers, and filling is done in cleanroom environments. Filling is done in an enclosed area where a stream of recirculating laminar-flow filtered air forms a screen preventing entry of microorganisms into the package during the filling process. High-efficiency particulate filters screen out microorganisms from the recirculating air. Proper worker attire and an overall low microbial load in the air near the filling equipment reduces the risk of microbial contamination of the product. Pharmaceutical preparations generally do not support microbial growth; therefore chance contamination of one viable microorganism in a package would not be a problem. However, contamination by just one viable pathogenic microorganism in a food that supports its growth will have disastrous results.

Testing. Aseptic packaging machines must be tested for their ability to maintain sterile filling conditions before they are placed into commercial operation. Typical tests include installing test strips containing indicator organisms inside the filling station to test the ability of the system to sterilize this section; and inoculating packaging material with a test organism and aseptically packaging the product to determine if the test organisms survived and spoil the product. Finally, the aseptic packaging system is used to package the food product at a speed used in commercial operations over a period of several hours. The packages are then incubated over several weeks and evaluated for signs of spoilage. Spoiled products are evidence of aseptic packaging failure.

Package designs. Aseptically packaged products are commercially available in brick-shaped packages that consist of a laminate of layers of polyethylene, foil, paper, and polyethylene; multilayered plastic film laminates in a protective cardboard box (bag-in-box); thermoformed plastic cups; and pouches of plastic film. A number of machines have been developed to produce special packages. In one type (**Fig. 49**), the machine receives roll stock of packaging material that consists of a laminate of paper, aluminum foil, and plastic and forms this into brick-shaped packages. In another type (**Fig. 50**), the machine receives paper/foil laminate blanks and forms these into gable-top containers. Packages may be in individual portion size, institutional size for food service applications, or large sizes for industrial reprocessing.

Romeo T. Toledo

Bibliography. American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE), *2002 ASHRAE Handbook: Refrigeration*, 2002; N. A. Anderson, *Instrumentation for Process Measurement and Control*, 3d ed., 1997; G. V. Barbosa-Cánovas et al., *Preservation of Foods with Pulsed Electric Fields*, 1999; P. A. Belter, E. L. Cussler, and W. S. Hu, *Bioseparation: Downstream Processing for Biotechnology*, 1988; A. L. Brody, *Controlled/Modified Atmosphere/Vacuum Packaging of Foods*, 1989; A. L. Brody and K. S. Marsh, *The Wiley Encyclopedia of Packaging Technology*, 2d ed., 1997; D. Claxton, L. Svarovsky, and M. Thew (eds.), *Hydrocyclones '96*, 1996; P. Cooper (ed.), *Pumping Machinery*, 1993; C. Doona et al. (eds.), *High Pressure Processing of Foods*, 2006; D. L. Downing (ed.), *A Complete Course in Canning and Related Processes: Fundamental Information on Canning*, 13th ed., 1996; P. F. Fellows, *Food Processing Technology*, 2d ed., 2000; F. J. Francis (ed.), *Wiley Encyclopedia of Food Science and Technology*, 4 vols., 1999; C. J. Geankoplis, *Transport Processes and Unit Operations*, 4th ed., 2003; S. A. Goldblith and W. W. Rothmayr, *Freeze Drying and Advanced Food Technology*, 1975; G. W. Gould (ed.), *New Methods of Food Preservation*, 1995; R. Guy, *Extrusion Cooking: Technology and Applications*, 2001; B. Hallstorm, C. Skjodebrand, and C. Tragardh, *Heat Transfer and Food Products*, 1988; R. Hayashi (ed.), *Trends in High Pressure Bioscience and Biotechnology*, 2002; D. R. Heldman and D. B. Lund (eds.), *Handbook of Food Engineering*, 2d ed., 2006; Y. H. Hui (ed.), *Dairy Science and Technology Handbook*, 3 vols., 1994; Y. H. Hui, *Food Plant Sanitation*, 2002; A. Ibarz, *Unit Operations in Food Engineering*, 2002; K. Kulp and J. G. Ponte, Jr. (eds.), *Handbook of Cereal Science and Technology*, 2d ed., 2000; W. L. McCabe, J. C. Smith, and P. Harriot, *Unit Operations in Chemical Engineering*, 7th ed., 2004; K. R. Osborne and W. Jenkins, *Plastic Films, Technology and Packaging Applications*, 1992; R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineer's Handbook*, 7th ed., 1997; M. N. Riaz, *Extruders in Food Applications*, 2000; S. K. Sastry and B. D. Cornelius, *Aseptic Processing of Foods Containing Solid Particulates*, 2002; G. Robertson, *Food Packaging: Principles and Practice*, 2d ed., 2005; H. G. Schwartzberg and M. A. Rao (eds.), *Biotechnology and Food Process Engineering*, 1990; P. G. Smith, *Introduction to Food Process Engineering*, 2003; A. J. Stapanoff, *Centrifugal and Axial Flow Pumps: Theory, Design, and Application*, 2d ed., 1993; R. T. Toledo, *Fundamentals of Food Process Engineering*, 3d ed., 2006; P. Walstra et al., *Dairy Science and Technology*, 2d ed., 2005.

Food fermentation

Production of food with the aid of microorganisms, which may be yeasts, molds, or bacteria. Well-known examples of such foods are bread, cheese, beer, wine, vinegar, some sausages, sauerkraut, yogurt, and cultured milk. This article deals with fermented foods produced and consumed in large amounts by millions of people in Africa and Asia. In Japan, for example, soy sauce (shoyu) is consumed at the rate of 11 quarts (10 liters) per capita per year, and annual production is close to 330,900 gallons (1,252,600 liters). Production of miso and natto is in the range of 625,900 and 174,200 tons (567,800 and 158,000 metric tons), respectively.

Fermented commodities include cereals such as wheat, rice, sorghum, corn; legumes, peanuts, soybeans, pulses; red meat, sausages, pork; milk; fish, shellfish; and plant juices. Fermented food is used as a staple, such as tempeh in Indonesia; as a condiment like soy sauce; as a coloring agent like the red rice used widely in the Far East; and as the breadlike product produced in India and the Near East. The food may be a liquid like the fermented fish sauces of Indochina; a paste such as miso made from rice and soybeans in Japan; or a solid, like tempeh, made by the mold fermentation of soybeans. *See* FERMENTATION.

Production. Fermented foods were developed in prehistoric times and are widely produced all over the world. Their continued use results from the mild biochemical changes produced by enzymes. For example, starches are broken down to the sugars required for a second fermentation. In soybean fermentations, protein yields amino acids, and lipases act on the oil. Fermentations produce vitamins and simple chemicals such as lactic acid; improve digestibility, odor, and flavor; change physical properties; and prevent spoilage. *See* PROTEIN METABOLISM.

Among disadvantages of fermentation is the possibility of food poisoning due to bacteria or molds. Also, some fermentations are slow (taking up to 8–12 months), and some product is always lost. *See* FOOD MANUFACTURING.

Microorganisms. Among the molds used in food fermentations are *Aspergillus*, *Rhizopus*, *Penicillium*, *Neurospora*, *Actinomucor*, *Mucor*, *Amylomyces*, and *Monascus*. The yeasts include species of *Saccharomycopsis, Zygosaccharomyces*, and *Candida*. The bacteria are *Bacillus* and lactic acid bacteria, including *Lactobacillus*, *Streptococcus*, *Pediococcus*, and *Leuconostoc*. *See* YEAST.

Inoculum. The inoculum or fermentation starter may be of four types. The substrate may be moistened, heat-sterilized, and inoculated with a single organism. In the commercial production of red rice, a single species of *Monascus* is used to inoculate rice, which eventually takes on a reddish-purple hue that imparts color to meat and wine. In a second type, more than one strain of a single species is used. Koji is the mold preparation containing more than one strain of *Aspergillus oryzae* that produces enzymes on rice, wheat, or soybeans. It is used in a second fermentation to make various types of miso and shoyu. A third type of inoculum contains more than one species of microorganism. Ragi is an Indonesian culture used in various fermentations of rice or cassava. Known by different names in other countries, the dry powder contains *Amylomyces*, *Mucor*, *Rhizopus*, yeast, and bacteria. Finally, some substrates contain a complex inoculum in which many different microorganisms of unknown identity are present. An example is nuocmam, a fish sauce. Small fish are packed with salt in earthen containers and allowed to ferment at room temperature for several months. The resulting liquid is removed and used as a condiment much like soy sauce. The salt prevents the development of food-poisoning microorganisms.

Applications. There are countless varieties of food fermentations in the world, sometimes known under different names in different countries. Of the following two examples, one is quick and simple, and the other slow and complicated.

Tempeh. Tempeh is made by a simple mold fermentation, which is being introduced in the West. Only soybeans are used as substrate. Dry beans are washed and soaked overnight at about 77°F (25°C). The seed coats are removed, and the beans are boiled for 30 min, drained, cooled, and briefly dried. After inoculation with spores of a *Rhizopus* strain, the beans are incubated for 20–24 h at 88–89°F (31–32°C) in plastic bags perforated for aeration. Acid may be added to control bacterial contamination. The fermentation produces a solid cake in which the beans are bound together with mycelium. The product, after being sliced into thin pieces and fried in deep fat, is brown and has a nut flavor.

Miso. Miso fermentation is a complicated two-step process. The product is used in seasoning and as a base for soups in the Orient. Polished rice is washed, soaked for 16 h, drained, steamed, cooled, and inoculated with the mold *Aspergillus oryzae*. After about 2 days at 82°F (28°C) with frequent turning, the moldy rice is called koji. Soybeans that have been crushed, soaked in water, drained, steamed, and cooled are mixed with koji, followed by the addition of 2 kg (4.4 lb) of salt for 15 kg (33.1 lb) of the mixture and an inoculum of special yeast and bacterial cultures. The mixture is allowed to ferment for 3 months or more and is mashed to a paste for sale. There are many kinds of miso, depending on the amounts of substrates and salt and the time and temperature of fermentation. *See* SOYBEAN.

C. W. Hesseltine

Bibliography. C. W. Bamforth, *Food, Fermentation and Micro-organisms*, 2005; L. R. Beuchat, *Food and Beverage Mycology*, 2d ed., 1987; Y. H. Hui, *Handbook of Food and Beverage Fermentation Technology*, 2004; G. G. Khachatourians and D. K. Arora (eds.), *Applied Mycology and Biotechnology:* vol. 2: *Agriculture and Food Production*, 2002; W. Shurtleff and A. Aoygai, *The Book of Miso*, rev. ed., 2001; K. Steinkraus (ed.), *Handbook of Indigenous Fermented Foods*, 2d ed., 1995; M. J. Waites et al. (eds.), *Industrial Microbiology: An Introduction*, 2001.

Food manufacturing

A total sequence of food operations, including the growth and selection of raw materials, harvesting, processing, preservation, and distribution. In general, the aim of all food manufacturing operations is to extend the availability of seasonal crops to year-round use.

Many of the technological operations involved in food manufacturing have a long history of use, beginning with the invention of fire and the cooking of foods, through the first large-scale food production industries in ancient Egypt and Rome, where bakeries provided consumers with bread. This was the beginning of manufactured convenience foods and, in addition, relieved consumers of the burden of making their own food products.

The products of food manufacturing differ from traditional foods of plant or animal origin which have undergone minimal treatment. For example, the quality of apples sold in the winter can be maintained, through the use of controlled-atmosphere storage, which retards the ripening process by controlling the levels of oxygen, nitrogen, and carbon dioxide in the atmosphere of the storage facility. Atmosphere control is also used to hasten ripening so that fruits may be harvested in the unripe stage for ease of handling and then ripened rapidly in storage. In other cases, the package itself allows the diffusion of only certain atmospheric gases and thus maintains quality. However, minimal treatment is not always directly correlated to quality and nutritional value. Foods that are harvested at optimum maturity, rushed to the manufacturing plant, cleaned, washed, cut, sorted, and processed by such means as rapid freezing are often higher in quality and nutrition than foods picked fresh and then stored under less than optimum conditions. Also, there are certain foods that cannot be maintained in a state close to the raw product. Tomatoes, for example, are not amenable to freezing or long periods of storage, yet they represent a major crop that must be harvested within a short period of time. If it were not for processing, tomatoes would not be available for human consumption throughout the year. Therefore, such products as heat-processed sauces, pastes, and stewed tomatoes have been developed. Other food products are even further removed from the raw product: oil is being produced from seed; and plant proteins are being used as extenders or substitutes for meat, as additives for nutritious beverages, and as bases for many formulated foods.

There are many other forms of food preservation representing both ancient and space-age technologies. The ancient operation of sundrying was first employed when it was realized that dried fruits remained wholesome and edible for long periods of time. Today, with the additional knowledge that drying, evaporation, and concentration all reduce the water activity or increase the osmotic pressure of a food to the point where bacteria will not grow, this technology is used for sophisticated products such as powdered milk and freeze-dried mushrooms. Food additives, such as salt, sugar, and other solutes, which reduce the water activity or increase the osmotic pressure, and acids, which inhibit bacterial growth, also achieve the preservation effect. Many food additives are natural in origin, and their preservative effect was noted in nature prior to their use as food additives. Freezing, heat sterilization (canning), pasteurization, fermentation, baking, and meat curing are other well-known forms of preservation. Irradiation processes for food have also been developed, and low-level irradiation has been approved in the United States by the Food and Drug Administration (FDA).

Food manufacturing, however, is not solely involved with the preservation of food but is also concerned with the production of high-quality, appealing, and wholesome food. To fulfill these goals, five broad categories of food additives are often used: flavors, coloring agents, preservatives, texturizing agents, and miscellaneous. The last category includes a variety of substances that may retain moisture, control acidity, act as leavening agents, or provide nutrients such as vitamins and minerals.

The final operation in the manufacturing process is that of packaging, which is governed by the physicochemical attributes of the food, the preservation process involved, the gaseous permeability desired, the conditions under which the product is to be stored, the desirability of viewing the product through a clear film or glass, and the expense.

Historically, metal and glass have been used to package heat-processed foods; more flexible films are used for foods which undergo less vigorous treatment. Adoption of the regulation allowing the use of hydrogen peroxide as a package sterilant has permitted the use of nonrigid flexible packages for heat-sterilized foods (aseptic packaging). This type of packaging is very cost-effective. *See* FOOD ENGINEERING; FOOD PRESERVATION. Fergus M. Clydesdale

Food Additives

Generally, a food additive is any substance used in small amounts in or on food to achieve a particular result. The terms food, ingredient, and additive are neither clearly bounded nor exclusive, and use determines how a substance is termed. Sugar in a sugar mint is a food; in a cake, it is an ingredient; in bread dough, it is a yeast food, an additive.

Some additives are direct (or intentional); that is, they are added to food to restore, enhance, or add desirable characteristics such as safety, stability, appeal, and economy, or to improve the effectiveness of processing. Others, called indirect additives, occur in food, usually in trace amounts, not because they confer any desirable property to the food itself, but because they inevitably are left as a result of migration from packaging materials, or from use in agriculture, storage, or transportation (**Table 1**).Certain direct additives may be regarded as second-order additives if they performed a useful function in some ingredient prior to its addition to a final food but lost that effective role in the final food. Thus, anticaking agents are often used in salt. If the salt

TABLE 1. Approximate numbers of direct and indirect food additives

Additive	Number
Direct	
Prior sanctioned	≈100
GRAS only	1558
Substances that are both GRAS and regulated food additives	868
Regulated food additives	312
Colors and coloring adjuvants	58
Subtotal	2896
Indirect	
Packaging, GRAS	110
Regulated	≈3000
Pesticides	300[*]
Subtotal	3410
Total	6306[†]

[*] Active ingredients only; there are thousands of different commercial formulations.

[†] An overestimate; an accurate count is impractical because of duplicate listings and difficulties of classification.

was used in making mayonnaise, the anticaking agent would no longer be useful and would then be a secondary additive.

Types. The daily average composition of the American diet consists of a number of categories (**Fig. 1**). There is a certain amount of overlap, because many substances fall in more than one category. In the food manufacturing industry, dietary consumption is referred to as disappearance. However, actual consumption is less than disappearance because of loss, waste, and nonconsumptive uses such as fermentation and processing. The category "other natural constituents" (Fig. 1) is numerically enormous—probably numbering in the hundreds of thousands—and includes essential oils, resins, alkaloids, nucleotides, phospholipids, enzymes, pigments, and hormones. Some of these substances, particularly certain essential oils and their components, are used in food manufacturing as direct additives. Most of the food additives comprise major ingredients that are generally recognized as safe (GRAS), including substances such as yeast, baking powder ingredients, and common spices. Many are carbohydrates (such as invert sugar), fats (mono- and diglycerides), or proteins (albumin). The distribution of use of direct additives (**Fig. 2**) ranges from sucrose at about 44 kg (97 lb) per person per year to some flavors and micronutrients at less than 0.1 nanogram per person per year. The median direct additive is consumed at about a milligram per person per year.

Technical effects. The purpose of all direct food additives is their physical or other technical effects (**Table 2**).

Other specific technical effects can be defined, most of them related to processing, such as freezing agents, and washing and peeling aids. In addition, there are many substances that serve several technical effects. Ascorbic acid (vitamin C), for example, is a nutrient supplement, an antioxidant, or a curing adjuvant, depending on the conditions of use. Also, the boundaries between the technical

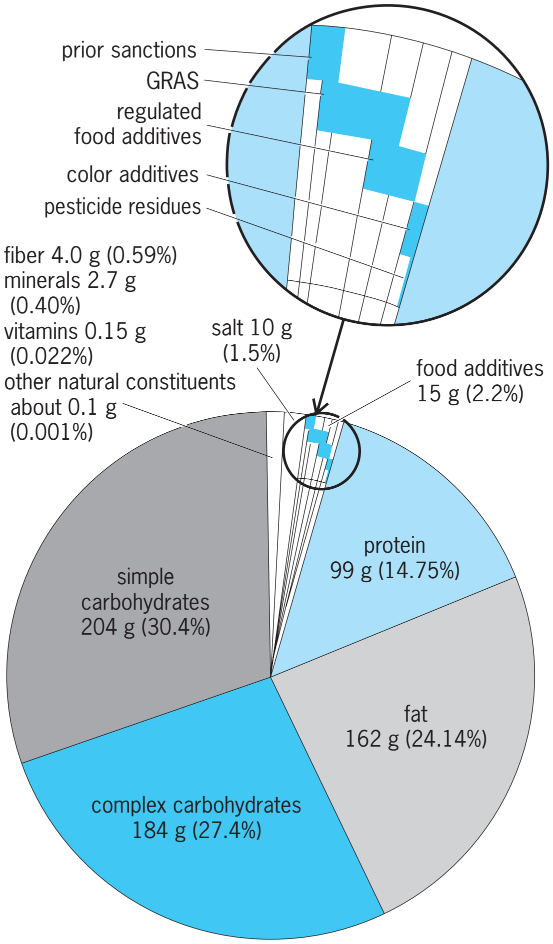


Fig. 1. Average composition of the United States diet dry weight per capita per day. Prior sanctioned substances were specifically approved by FDA or USDA for use in food prior to the passage of the Food Additives Amendment in 1958.

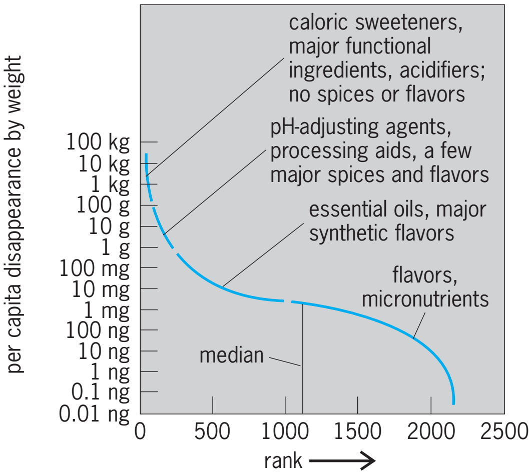


Fig. 2. Rank ordering of food ingredients by per capita annual disappearance.

TABLE 2. Technical effects of direct food additives

Category effect	Agent
Anticaking agents: prevent caking, lumping	Calcium stearate, cornstarch, sodium aluminosilicate
Antimicrobial agents (preservatives): prevent growth of microorganisms and consequent spoilage	Sodium benzoate, calcium propionate, potassium sorbate
Antioxidants: retard browning and rancidity	BHA, BHT, ascorbic acid (vitamin C), ethoxyquin
Colors and coloring adjuncts: impart, preserve, or enhance color	FD&C Blue No. 1 and other certified synthetic colors, β-carotene and other uncertified synthetic colors, beet powder, caramel and other natural colors
Curing and pickling agents: impart flavor and color to cured meats, discourage spoilage, prevent bacterial toxin formation	Sodium nitrite, salt, sodium metaphosphate
Dough conditioners, strengtheners: make dough easier handle, bread more appealing to consume	Potassium bromate, acetone peroxide, glyceryl monostearate
Drying agents: absorb moisture	Specially dried cornstarch, anhydrous dextrose
Emulsifiers: prevent separation of oil and water mixtures	Mono- and diglycerides, lecithin, sorbitan monostearate
Enzymes: promote reactions	Rennet for producing cheese curd, papain for tenderizing meat, pectinase for clarifying beverages
Fermentation and malting aids: promote rapid and proper fermentation	Gibberellic acid, potassium bromate
Firming agents: enhance crispness of plant tissues	Calcium salts, aluminum sulfate, calcium lactobionate
Flavor enhancers	Soy sauce, monosodium glutamate (MSG), disodium inosinate, disodium guanylate
Flavoring agents and adjuvants: natural and artificial flavors, fixatives, solvents	Spices, herbs, essential oils, and individual chemicals of both natural and synthetic origin
Flour-treating agents: bleach and mature flour	Acetone peroxide, benzoyl peroxide, potassium bromate
Formulation aids: carriers, binders, fillers, film formers, and so on	Starch, gum acacia, propylene glycol
Fumigants: kill undesirable organisms	Methyl bromide, ethylene oxide, phostoxin
Humectants: prevent drying	Sorbitol, propylene glycol, sodium tripolyphosphate
Leavening agents: impart light, fluffy texture	Yeast, monocalcium phosphate, sodium carbonate, and other baking powder ingredients
Lubricants and release agents	Oleic acid, hydrogenated sperm oil, mineral oil
Masticatory substances for chewing gum	Chicle, rubber, paraffin, glycerol esters of rosin
Nonnutritive sweeteners: sweeten with few or no calories	Saccharin, aspartame, acesulfame-K
Nutrient supplements: enhance or restore nutritional value	All the known essential nutrients such as vitamin A and other vitamins, iron and other trace minerals, amino acids and essential fatty acids
Nutritive sweeteners: sweetness with calories	Dextrose, fructose, sucrose
Oxidizing and reducing agents: remove or modify ingredients or contaminants that would damage quality	Glucose oxidase (enzyme) to destroy glucose in dried egg, so that it will store well; hydrogen peroxide added as a bleaching or antimicrobial agent
pH control agents: reduce or increase or stabilize acidity	Vinegar (acetic acid), sodium bicarbonate, hydrochloric acid
Processing aids: aid in clarifying, filtering, crystallizing, and so on	Charcoal, diatomaceous earth, tannic acid
Propellants, aerating agents, gases: expel a product, induce foaming, exclude oxygen	Nitrous oxide, carbon dioxide, nitrogen
Sequestrants: combine with trace metals to promote stability	Citric acid, ethylenediaminetetraacetic acid (EDTA), phosphoric acid
Solvents, vehicles: dissolve, suspend, or carry	Alcohol, propylene glycol, glycerine
Stabilizers, thickeners: increase viscosity, improve mouthfeel, prevent separation	Starch, modified food starches, carrageenan
Surface-active agents: promote rapid wetting and dispersing, enhance or prevent foam	Sodium lauryl sulfate, lactylic esters of fatty acids, dimethyl polysiloxane
Synergists: disproportionately enhance effects of other additives	Citric acid, tricalcium phosphate, and other phosphates
Texturizers: affect feel or appearance	Sodium bicarbonate, glycerine, corn syrup, modified food starch

effect categories are not always clear; that is, some flour-treating agents are oxidants, emulsifiers may act as surfactants, and some formulating and processing aids may act similarly.

While there are usually several substances available for achieving a particular technical effect, they rarely are fully interchangeable. For each application there is usually a single best choice if the requirements can be defined adequately.

Legal definition and approval for use. The definition of food additive in the Federal Food, Drug, and Cosmetic Act (FD&C Act; 21 U.S. Code 321) is far more complex than the practical definition given above. Section 201 first broadly defines the term to include all components of food, then makes exclusions: "The term 'food additive' means any substance the intended use of which results or may reasonably be expected to result, directly or indirectly, in its becoming a component or otherwise affecting the characteristics of any food . . . if such substance is not generally recognized, among experts qualified by scientific training and experience to evaluate its safety, . . . to be safe under the conditions of its intended use; except that such term does not include . . . a pesticide chemical . . . or . . . a color additive; or any substance used in accordance with a sanction or approval granted prior to the enactment of this paragraph . . . ; or a new animal drug."

The words "generally recognized . . . to be safe" have been lifted out of the language of the act, paraphrased as "generally recognized as safe," and reduced to the acronym GRAS. Thus, GRAS substances, such as salt or pepper or vanillin (a flavor), though in a practical sense food additives, are excluded from the legal definition of food additive. Also excluded are pesticides and fumigants, colors such as beet juice or FD&C Yellow No. 5, and "prior sanctioned" materials such as sodium nitrite used in curing meats.

In Section 409, the act requires any food additive, so defined after exclusions, to be regarded as "unsafe" unless it is either exempted for investigational use, or is used following a regulation prescribing conditions for its safe use.

The information required and the specific steps involved in approval of a new food additive, in the common sense of the term, depend on the particular legal classification of the substance, although the pattern of information in all cases is generally similar.

For a substance that is legally a food additive, a petition must be filed with the FDA for a regulation permitting its use. Section 409 of the FD&C Act sets out the information required in such a petition. Included are the name, chemical composition and identity, the conditions of proposed use, the technical effect (purpose) of the additive and the amount required to produce that effect, analytical methods for determining the additive in food, and reports of all safety testing. The FDA may also require manufacturing details and samples. Neither the statute nor regulations spell out how much toxicological and related information of safety will be required. The FDA has based its requirements in each instance on an ad hoc determination of the kinds of data required to provide assurance of safe use. Expected intake, the results of previous testing, and accumulated knowledge of the toxicity of similar materials all weigh heavily. These judgments, made over years, have been collected and collated into a compilation for providing flexible general guidance called the "Red Book," published by the U.S. Department of Health and Human Services.

If the firm desiring to use the substance believes that it would be more appropriate to regard the substance as GRAS, it may petition the FDA to affirm that judgment. The data requirements are similar to those for any food additive.

The FD&C Act, however, does not define which experts shall conclude that a substance is of GRAS status. It merely requires that they be "qualified by scientific training and experience to evaluate its safety." Thus, the experts may be drawn from either the public or private sectors. Although a private determination that a substance is generally recognized as safe may seem contradictory, it is conceivable that a firm could reach this conclusion on its own. On a broader basis, the flavor industry has regularly published the judgments of a panel of independent experts on general recognition of safety, and most new flavoring substances have come into use by that route. Again, the data and judgments required are generally similar to those for all food additives.

Color additives are approved by a similar process under Section 706 of the FD&C Act, except that there is no provision for GRAS status. "Prior Sanctions" apply only to certain substances approved for use before 1958 and are no longer granted.

Pesticide residues on food are approved under Section 408 of the act, with generally similar provisions. A pesticide, however, must first be registered under the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA), and the Secretary of Agriculture must also certify the pesticide's usefulness for the purposes proposed.

The Food Additives Amendment and Color Additive Amendments contain the so-called Delaney Clause. This denies regulatory approval to a substance "found to induce cancer when ingested by man or animal. . . . " With the passage of time, analytical methods have become ever more sensitive in finding minute traces of contaminants and impurities. More comprehensive and severe conditions of toxicological testing have raised both questions of possible carcinogenic potential in many materials tested and questions about the relevance of the tests to human safety. These trends have increased debate over providing more flexibility, through either administrative interpretation or legislative change, than the clause now appears to allow. Such new knowledge caused Congress to provide limited exceptions to the clause for animal drugs and nutrients. The clause does not apply to GRAS and prior sanctioned substances. *See* TOXICOLOGY. Richard L. Hall

Food Analogs

The terms food analog, structured food, fabricated food, texturized food, imitation food, and architectural food are all used to describe foods which are considered to have equal or superior characteristics to those of the foods which they are designed to imitate. These enhanced characteristics usually involve sensory properties, better quality control, superior nutritional properties, or manufacturing economies.

Food analogs usually require a high degree of design and engineering because the food product should have unique functional and marketable properties which can be distinguished by the consumer. Food analogs are usually based on soy, and include fabricated meats, seafoods, cheese, vegetables, fruits, and eggs. They have been available since about 1960 as novelty products, but have gained consumer acceptance as meat substitutes for vegetarians and individuals with special dietary restrictions, generally because they are cheaper than meat. In addition, they are used extensively in food service programs, especially to extend hamburger. *See* SOYBEAN.

Technologies. Once the structuring mechanism or process is chosen for food analogs, customized manufacturing equipment and machinery for high-speed production conditions must be designed and constructed.

Fiber formation. Fiber formation technology is based on the chemistry and technology of proteins, particularly soy proteins. Soybeans are hulled, and the oil is removed by solvent extraction. Soybean meal contains factors that must be inactivated by moist heat. For food uses, the processing may consist of simple heating and grinding of the defatted material as in the preparation of flours and grits, or of further fractionation to increase protein content as in the production of concentrates and isolates. *See* SOLVENT EXTRACTION.

Soybean protein concentrates consist of a minimum of 70% protein plus the polysaccharides, and are made by three methods that render the major proteins insoluble while the low-molecular-weight compounds are removed. One process consists of extraction with aqueous alcohol. A second process involves extraction with dilute acid at pH 4.5 (the isoelectric point and region of insolubility of the major proteins); the acid-leached concentrate is neutralized before drying. In a third process the proteins are rendered insoluble by heat denaturation; the low-molecular-weight constituents are then washed out with water. *See* ISOELECTRIC POINT; PROTEIN.

Protein isolates are the most refined soy proteins available. They are prepared by removing all water-insoluble polysaccharides, as well as water-soluble sugars and other minor constituents. Defatted flakes and flours of high-protein solubility are extracted with dilute alkali at 120–130°F (50–55°C). After the insoluble residue (water-insoluble polysaccharides plus residual protein) is separated by screening, filtering, and centrifuging, the extract is adjusted to pH 4.5 with food-grade acid. When the major proteins are brought to their isoelectric point, they precipitate. This protein curd is filtered or centrifuged from the solubles and washed. The curd may be neutralized and spray-dried.

Isolates are the basic ingredient of the spun-fiber meat analogs. In this process, the soy protein is resolubilized in an alkali aqueous media. This liquid may be extruded through tiny holes and then recoagulated in an acid bath in the form of fibers in a manner similar to making rayon. The fibers can then be spun into ropes, with the final texture approaching the fibrous structure of chicken or beef muscle tissue. The fabricated tissue can be interlaced with fats, food flavorings, and food colors. The products also may be dehydrated, compressed, or otherwise processed. This process was first adapted from the production of synthetic textile fibers, such as rayon and nylon (**Fig. 3**).

Because the manufacturing steps are expensive, the soy meal may be processed instead directly in a heated extruder to form an expanded cooked soy matrix which can be incorporated into other foods.

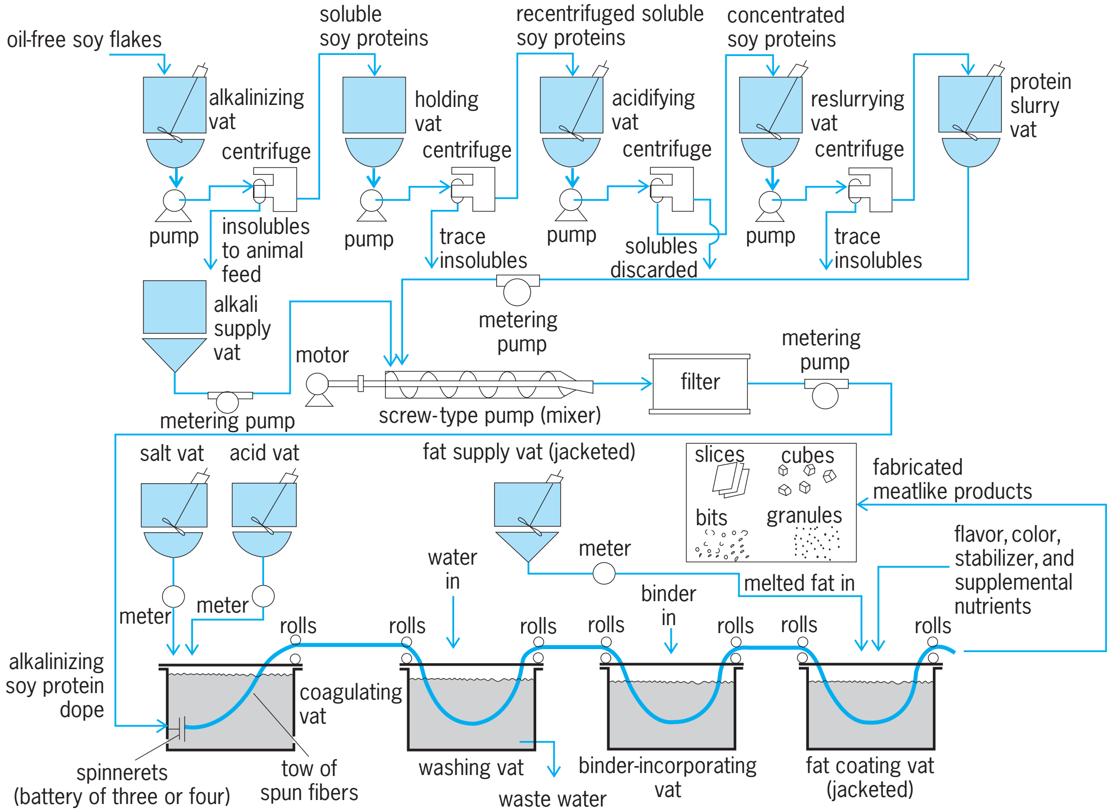


Fig. 3. Flow chart of the fiber-spinning technique used to prepare food analogs. (*After N. R. Lockmiller, Increased utilization of protein in foods, Cereal Sci. Today, 18*(3)*:77–81, 1973*)

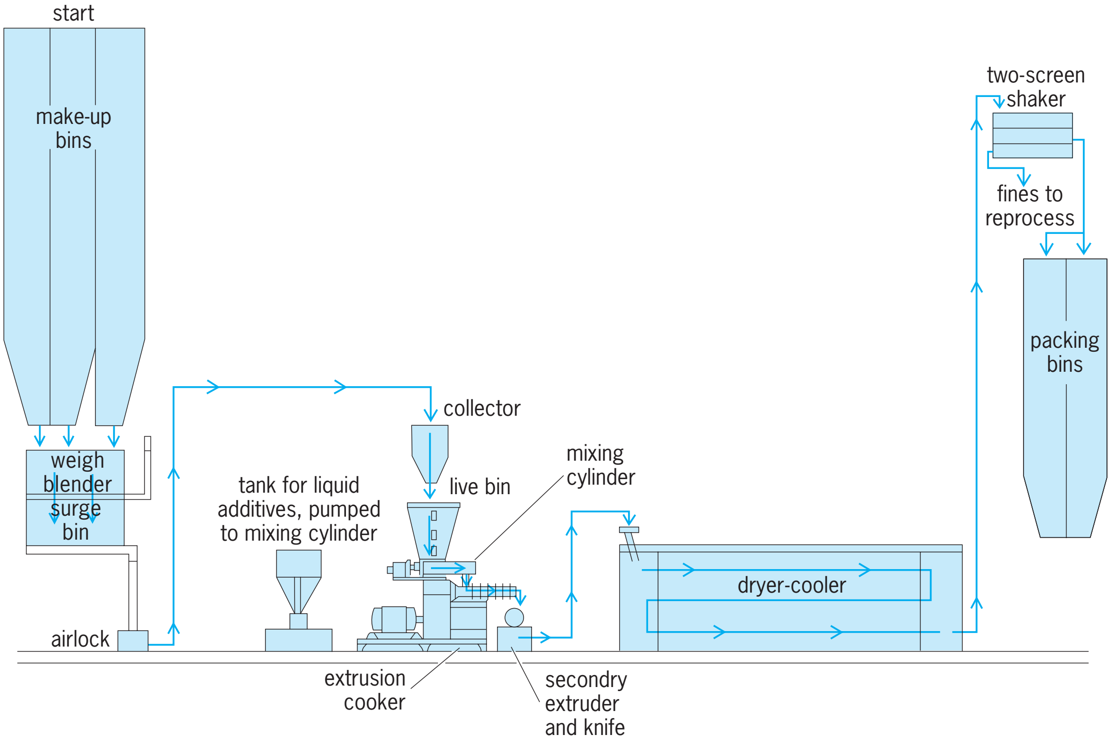


Fig. 4. **Typical high-temperature extrusion process for production of a meat or seafood analog. (*After O. B. Smith, in L. D. Hill, ed., Textured Vegetable Proteins, World Soybean Research, Interstate Printers and Publishers, 1976*)**

Extrusion. There are two basic types of extrusion technology, high-temperature and low-temperature. High-temperature extrusion is associated with a cooking process where a number of chemical changes take place in the dough or matrix (**Fig. 4**). Examples of these are starch formed into granules, gelatinized proteins becoming thermoplastic gels, inactivation of the antitrypsin factor in soya, and soluble proteins being denatured. Cold-temperature extrusion is a process that is dependent on compression rather than on chemical reactions. With dies designed especially for the process, many sophisticated analogs can be developed by simple combination of the raw ingredient, binders, flavors, and suitable additives.

Chemical setting. Internal setting by chemical reaction is used to prepare fruit analogs that have fairly uniform texture such as apple, peach, pear, and apricot. A two-feed process involving rapid mixing is employed. One feed, which is near pH 7, contains alginate and the calcium-ion source anhydrous dicalcium phosphate; the other feed contains fruit puree, sequestrant, and acid. Since the alginate is predissolved, no alginate hydration problems are encountered. Dicalcium phosphate is insoluble at neutral pH. The structured fruit is prepared by pumping the two mixes through a suitable high-speed mixer and then allowing the final mixture to set under shear-free conditions. The gelling reaction is brought about by the calcium ions released from the dicalcium phosphate, which dissolves as the pH is lowered on contact with the acidic puree phase. The product can next be extruded onto a slab on a moving conveyor and then be diced or cut into uniform shapes prior to further processing. This method is also used to produce pimento pieces.

There are also two other chemical processes which employ alginates—diffusion setting and setting by cooling—and these also are used in the fruit industry to produce structured fruits. *See* ALGINATE.

Types. Analog foods may be divided into three groups: dairy products, meats, seafood, and egg products; fruit and vegetable products; and farinaceous products.

Dairy products. Margarine was one of the first food analogs to be introduced into the marketplace in the United States. This term applies to certain types of shortenings as well as table spreads. The total consumption of table spreads in the United States has remained rather constant since 1970, but butter has declined in popularity while margarine has increased, mostly because of concerns over heart disease and cholesterol.

In the United States, margarine is made largely from vegetable oils that have been hydrogenated and crystallized to remove high-melting triglycerides in order to achieve the proper spreading texture. The vegetable oils may also be blended with lesser quantities of animal fats. In the United States, margarine must contain no less than 80% fat by law. To provide proper spreadability and reduce calorie consumption, whipped butters and whipped margarines are made by incorporating more water and air into them. *See* MARGARINE.

Imitation milks made from soy protein isolates and vegetable oil have met with limited consumer acceptance. Improvement in quality, plus a greater economic advantage over cow's milk, seems necessary before these products become successful. These products have become available in paper-foil containers as a result of the development of aseptic processing of liquids. *See* MILK.

Frozen dairy novelties with artificial fruit flavors have been marketed for many years, targeted mainly for children. Structured frozen pudding products on a stick have been manufactured as well. This product is manufactured by using freezing equipment and a formula that will minimize dripping while allowing excellent textural and other sensory properties.

Many cheese analogs with a variety of characteristics are available. The key benefits of cheese analogs are linked to economics and nutrition, while some products are formulated to have improved melting qualities, such as on pizza. These analogs are prepared by comminuting and mixing, with the aid of heat, one or more types of cheese plus soy protein and certain emulsifying salts; water, seasoning, and color may be added. When certain optional dairy ingredients such as cream, skim milk, whey, or their solids are added to the cheese blend and processed, the product is called a pasteurized process cheese food or a pasteurized process cheese spread. Both products may contain less fat and more moisture than process cheese as prescribed in the Federal Definitions and Standard of Identity. In addition, the pasteurized process cheese is spreadable at 70°F (21°C). Cheese analogs are less subject to spoilage and shrinkage in weight than natural cheeses. Advantages to the consumer are a more uniform flavor, no waste, and excellent keeping quality. Structured cheese bits are also available. They have good microbiological stability and integrity and are used for incorporation into other food products, such as meats and salad dressings. *See* CHEESE.

Meats, seafood, and eggs. The use of soy proteins to manufacture simulated meats became very popular in the 1960s and was enhanced by rising beef prices. However, simulated meats (meat, poultry, and seafood analogs) have not achieved the popularity that was predicted. While extruded soy products have a cost advantage over the spun-fiber items, they contain residual flavors plus the oligosaccharides stachyose and raffinose, which appear to cause flatus when soy is ingested.

Bacon-flavored textured soy products are sold in retail stores and are marketed to hotels, restaurants, and institutions. Beef- and chickenlike products have been developed for addition to canned chili, dry-mix chicken ala king, and Spanish rice. Other uses are as replacement for part of the freeze-dried meats in dry soup mixes and for simulation of meat particles in dips, crackers, and snacks. Fine particles of textured soy flour are used as extenders in beef and pork patties, where they reduce the stickiness of meat, thereby permitting better release from patty-making equipment.

Much of the meat analog production is designed for the food service industry—hotels, restaurants, armed services, and institutions—which makes up a sizable segment of the food market. The scarcity of skilled workers in the food service industry places a premium on convenience in food preparation. Prepared foods containing meat analogs have been developed that require only thawing or reconstitution and heating.

A striplike bacon analog has been on the market for a number of years. This product is made by laying down spun isolate fibers randomly and holding them together with an edible binder. Alternating layers of red and uncolored fibers are employed to simulate the lean and fat portion of bacon. The product is shaped, set by heat, sliced, and finally frozen. For consumption, the product is merely heated. Since cooking or frying is not necessary, there is no shrinkage, as compared with regular bacon which shrinks to 25% of its original weight.

Other innovative meat analogs are the formed meats. The product is formed by compressing meat under high pressure to simulate various structures, for example, ribs or patties. The benefits are portion control and a lean and boneless product.

Examples of fabricated seafood products are the so-called surimi (compressed fish and shellfish mixtures with added flavor and binding agents) and fish rings. Fish rings of uniform size and texture can be fabricated under high-speed conditions at refrigeration or freezing temperature. This method is used for products from clam, squid, crab, shrimp, scallop, and abalone.

The first egg analogs were products that imitate scrambled eggs when cooked, but do not contain cholesterol. These products are made from various egg proteins and dairy ingredients with added flavor, color, emulsifiers, and stabilizers. The ingredients are blended together, pasteurized, homogenized, and packaged for refrigerated storage.

Fruits and vegetables. Popularity of structured fruits has grown because of uniformity in regard to quality, color, size, and taste. These products are also low in moisture content, convenient, and steady in supply because of independence of seasonal harvest. In the manufacturing of these products, stabilizers such as pectin, locust bean gum, and xanthan gum are used.

One of the largest applications for structured vegetables is in Europe, where pimento strips are manufactured for use in stuffed olives. Structured pimento strips made from alginate and guar gum allow for automatic stuffing at a reduced cost, compared to regular pimentos, which are fragile, require manual stuffing, and have an uncertain supply position.

The manufacturing process can be adapted to produce such products as structured tomatoes, beets, carrots, zucchini, green peppers, and mushrooms. The process can also be adapted to produce dehydrated vegetable pieces, such as onions, peppers, carrots, and peas, for such applications as dry soups and sauces, where uniformity and rehydration properties are important. Structured onion rings are widely used in the food service industry because they provide excellent portion control, whereas natural onion rings vary in size, leading to serving problems as well as complicating batter and breading operations. Structured onion rings are made by producing a mash from dehydrated minced onions, flour, salt, and alginates and forming it into rings by a machine similar to that used to produce doughnuts. The rings are dropped into a calcium chloride bath where they acquire alginate skins, giving the rings sufficient rigidity to undergo the breading and frying process.

Farinaceous foods. The various possibilities for making farinaceous foods available through extrusion and compression are enormous. Products include granola bars, corn chips, potato snacks, and breakfast cereals containing soy flour as an additive to improve the amino acid balance of cereal proteins and to increase protein content. M. K. Schmidl; T. P. Labuza

Bakery Products

This category of foodstuffs includes such an extraordinarily diverse group of products that it is difficult to define. It is generally agreed that baked products contain, as a minimum, flour, water, and salt, but there are even exceptions to this rule—for example, flourless "cakes" and salt-free bread. Although the name indicates that they are cooked by baking in an oven, many varieties having similar composition, uses, and appearance are fried, for example, doughnuts and certain types of breads.

Formulations. There are numerous formulations for producing bakery products, providing a wide variety of nutritional choices.

Bread and rolls. Acceptable white bread can be made from flour, water, salt, and yeast. Italian bread is usually based on this simple combination of ingredients, and French and Vienna breads are seldom much more complicated. These basic breads tend to have a hard crust that is light in color, a coarse and rather tough interior, and a flavor that is excellent when the bread is fresh but considerably less desirable when it has been out of the oven several hours. Sweeteners, milk, and shortening are often added to improve flavor and texture; and eggs are used occasionally in specialty loaves. Barley malt syrup is very helpful in controlling fermentation. Sometimes a "sour dough" or starter consisting of a piece of dough from a previous batch that has been held a few days is used instead of, or to supplement, the yeast. Mold inhibitors and starch complexing agents (such as monoglycerides of fatty acids) can be used to improve the storage life of the bread. Yeast foods, enzymes, and emulsifiers are sometimes included to modify the processing response of the dough. Most bread is made with enriched flour, which contains several added vitamins and minerals.

Loaves are generally regarded as bread units which are more than about 8 oz (0.2 kg) in weight and are intended to be cut into chunks or slices for individual service. Rolls are considered to be portion-sized pieces, and they often have decorative conformations (as Kaiser rolls) and may be topped, as with sesame seeds, onion pieces, or poppy seeds. Both loaves and rolls may be "washed" with a fluid made from combinations of eggs, milk, and sugar before baking in order to enhance appearance of the crust. Hot rolls may be brushed with melted butter or other shortening to give a shiny crust and an attractive aroma.

There are many variations in shape, flavor, texture, and adjuncts (such as toppings and fillings). Mixtures similar in composition to those described above can serve as the basis for pizza crusts, pita bread, soda crackers, English muffins, and bagels by varying the postmixer processing steps. Variations in composition to include whole-wheat flour, and meals from rye, corn, and oats, as well as many noncereal ingredients such as potato flour and soy derivatives, yield a number of familiar items.

Sweet yeast-raised products. By increasing the levels of shortening, milk, and sweeteners in a basic bread formula, many products suitable for consumption as snacks or desserts can be prepared. These items are generally spiced and flavored and may also include adjuncts such as fruit gels, nuts, and pastry creams. Doughnuts, sweet rolls, coffee cakes, and Danish pastry are examples. Traditional Danish pastry is made from a rich dough which is sheeted out and layered with butter, after which it is repeatedly folded and sheeted until it is "made up" into fancy rolls or coffee cakes. The repeated folding and sheeting give Danish pastries a fine, silky texture with a rather flaky crust.

Chemically leavened products. This exceedingly diverse category includes all those bakery foods which rely on baking soda (sodium bicarbonate) as the source of carbon dioxide. Soda is usually added in the form of baking powder, which contains baking soda, some inert materials such as starch, and an acidic substance that will react with the soda to generate carbon dioxide. If baking soda is added separately to the mix, it is necessary to put some acidic ingredient such as molasses, sour milk, or cream of tartar into the batter. Chemically leavened products include cakes, cookies, certain kinds of crackers, cake doughnuts, most types of muffins, soda biscuits, and pancakes. Generally, these products are slightly to moderately denser than bread and rolls, and they are more crumbly due to a less elastic internal structure. In the absence of added flavors, they are also blander than yeast-leavened products, because they lack yeast and the by-products of fermentation.

On a flour-weight basis (that is, flour = 100%), the sugar content of batters for white or yellow layer cakes should fall in the range 110–160%. Shortening should be about 30–70%, while eggs should at least equal the added fats and oils. Liquid milk (whole or skim) should exceed the amount of sugar by about 25–35%. Soda should be added in an amount equal to about 1.2–2% of the flour weight, with salt 3–4%. Obviously, these general rules are often violated to secure special effects or to compensate for special processing limitations or ingredient problems.

Drop-type cookies can be made from cake batters, but commercial cookie formulas usually contain much less water and often contain less flour than the conventional cake batter. Some cookies, such as sugar wafers and shortbreads, contain no baking soda. The viscosity of the dough or batter must be adjusted to meet the requirements of cookie-forming machines, which are often very sensitive to small changes in the material to be processed.

Miscellaneous products. Angel food cakes and most pound cakes do not contain any leavening compound. They rely on the air whipped in during mixing for their porous structure and expansion in the oven. Pie crusts are usually made of simple mixtures of flour, fat, salt, and water; since expansion of the dough is not required, leaveners are not added. Turnovers, popovers, and cream puff and eclair shells are leavened by expansion of water vapor and air in their interior as their temperature rises during baking. These batters often include relatively large percentages of shortening or eggs.

Processing equipment and methods. Although both the individual pieces of equipment and the methods of assembling them into production lines differ greatly for different kinds of products, bakery equipment functions can be generalized as measuring, mixing, holding doughs in controlled environments, separating into pieces of predetermined weight, shaping or forming, baking, combining with adjuncts, cooling or tempering, and packaging. The sequence in which these steps are performed is not necessarily in this order, and steps may be omitted or added for specialized products.

Measuring and mixing. In large commercial bakeries, the measuring of ingredients is usually done automatically by meters and scales that receive bulk materials from tanks and silos. Minor ingredients such as colors and flavors are weighed and added manually in most plants. Mixing devices and processes have many features in common with their counterparts in other food factories, but two functions peculiar to bakeries—dough development and air incorporation—require special designs.

The term development is applied to a type of mixing that yields an elastic, extensible dough. Development is accomplished by using mixer agitators that press, stretch, and fold the dough while causing relatively minor amounts of tearing and cutting. Doughs for bread and rolls are typical examples of the intermediates requiring this type of mixing. A common type of mixer suitable for dough development has a large U-shaped container enclosing three or four thick cylindrical rollers which are oriented lengthwise in the trough and move in a circular pattern around the horizontal axis of the trough during the mixing cycle (**Fig. 5**).

Air incorporation occurs in virtually all doughs and batters as a normal and expected result of the mixing process. The size and distribution of the air bubbles have a pronounced effect on the internal structure and ultimate size of the finished product. Angel food cake batter is an example of a material that is leavened entirely by the air bubbles whipped into it during mixing, and special agitators are needed to fold in the air and subdivide it into small, uniform bubbles which are optimal for these cakes.

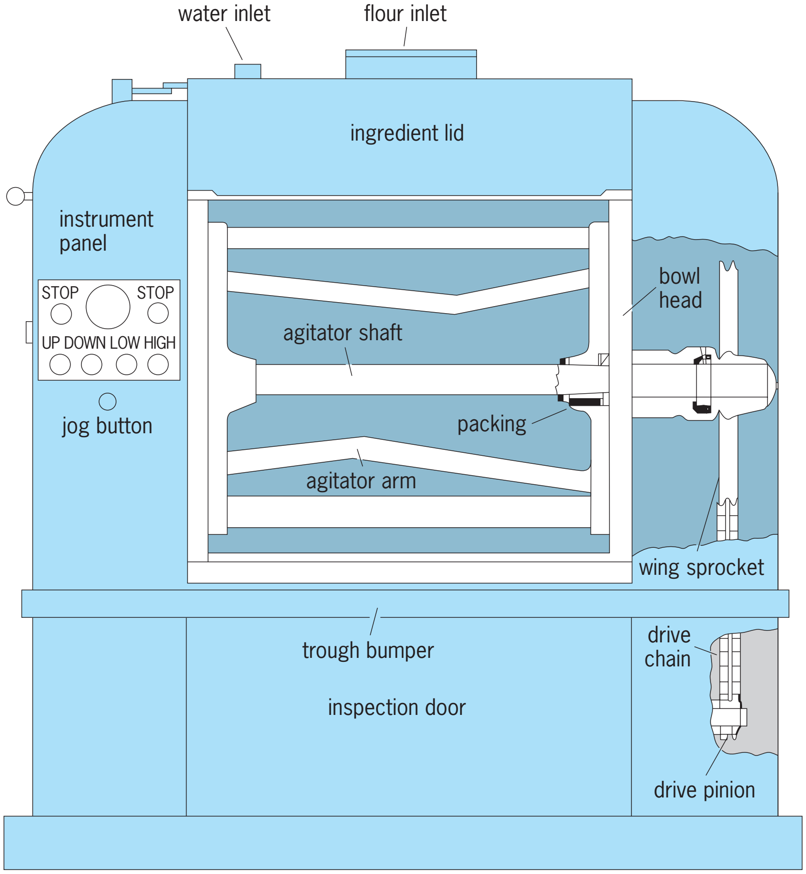


Fig. 5. Front view diagram of a horizontal dough mixer.

Fig. 6. Small integrated bread make-up system. A mass of bread dough is placed in the hopper of the divider (right) which cuts it into pieces of uniform weight. The rounder (center foreground) receives the rough pieces from the divider, forms them into spheroids, and delivers them to the overhead proofer. The dough undergoes a fermentation period while traveling through the proofer in individual pockets. The proofed dough pieces fall into the molder (center background) which sheets the dough and then curls it into a cylinder ready for depositing into baking pans. (*Adam Equipment Co.*)

Controlled environments. Virtually all yeast-leavened doughs must undergo one or more fermentation periods during which the yeast metabolizes carbohydrates and causes numerous changes in the dough. In order to control the rate of these reactions and maintain the dough in suitable condition for further processing, the dough must be held in rooms or cabinets that are maintained within narrow temperature and relative-humidity ranges. In its crudest possible form, the container will consist of a large tub or trough covered with a board. More sophisticated devices include conveyors with individual pockets for dough pieces traveling in cabinets provided with means for closely controlling the temperature and humidity. Temperatures are maintained by circulating air through steam-heated or electrically heated radiators, while high humidity is ensured by injecting water vapor into the cabinet. A high relative humidity is needed to prevent drying out, crusting, and loss of weight.

Dividing and rounding. These steps are used in processing yeast-leavened doughs. They are distinct from the final forming and shaping processes that establish the contours of the finished product. The mass of dough that comes out of a mixer (sometimes having undergone a prior fermentation stage) is dumped into the hopper of a divider, which squeezes a small amount of the dough into a chamber of adjustable volume (**Fig. 6**). The size of the dough piece is adjusted to give a baked product having the desired weight. Since the undivided dough mass continues to ferment and change in density, the divided dough pieces may slowly change in weight as the process continues, even though the divider presses much of the gas out of the dough.

From the divider, the dough pieces are carried immediately to a rounder, if bread loaves are being made. The rounder rolls the dough along a spiral guide fitted to the contours of a vertically oriented cone-shaped rotor. The cone is usually ridged or corrugated and rotates so as to mold the dough piece into a roughly spherical shape as it moves up the guide. The action develops a relatively smooth surface on the dough piece, thereby slowing down the loss of carbon dioxide and water during subsequent steps. The dough is then proofed, or held at controlled temperature and humidity for a length of time necessary to develop the desired handling characteristics and internal structure. The flavor of the finished product also depends to a considerable extent on the reactions occurring during the proofing step.

Forming the product. There are many kinds of dough-forming equipment in use of which the following are examples. Most cookies are formed in one of three ways: (1) dough is pressed into a die cavity engraved in the surface of a cylinder and then removed by suction; (2) soft, batterlike mixtures are extruded through shaped orifices and the strand cut into individual cookies by wires or pulled apart by the relative motion of the extruder and the oven band; or (3) doughs are sheeted out and stamped into individual circles, squares, or other shapes by reciprocating or rotating dies.

To form finished rolls and loaves of bread, various combinations of sheeting, cutting, curling, and

Fig. 7. Tunnel oven. Loaf-size dough pieces enter the oven at left background and are carried through the baking chamber on a continuous metal belt which serves as the hearth. The baked loaves exit at right foreground. Ordinary pan bread would require sets of pans conveyed by a slightly different mechanism, but the basic oven structure would be the same. (*Adam Equipment Co.*)

stamping are used. In some fully automatic bread production lines, a very soft cylinder of dough is extruded from the mixing chamber directly into the baking pan.

Baking. The baking process may take a few minutes for some crackers and cookies to well over an hour for large loaves of hearth breads. The usual type of oven used for the mass production of bread, cookies, and many other types of baked products is a long chamber, open at both ends, through which the products are carried on a continuous metal belt (**Fig. 7**). These tunnel ovens are commonly heated by gas flames burning inside the chamber, but some are heated by electricity, oil, or steam. The burners or other heating elements can be external to the main chamber, with hot air being carried into the oven by blowers.

Cooling or tempering. It is important to bring the temperature of a baked product to within a few degrees of normal room temperature before it is packaged. For small, low-moisture products such as crackers, cookies, and pretzels, cooling must be done slowly so that moisture remaining after the item leaves the oven has an opportunity to redistribute more or less uniformly before the piece hardens. Otherwise, the product may crack spontaneously during subsequent handling and storage. If large pieces of relatively high moisture content, such as loaves of bread, are packed while they are hot, moisture may condense on the crust, leading to more rapid spoilage from mold growth. Refrigeration is seldom used for cooling; most bakeries rely on the forced circulation of room temperature air to perform this function.

Adjuncts. Toppings, fillings, icings, and other adjuncts may be added before or after baking, depending on their characteristics and the effect desired. In most cases, the applicators are fairly simple machines which sprinkle seeds, granules, spices, or other condiments as the rolls pass beneath a hopper equipped with a rotating distributor bar at the bottom.

Packaging. Packaging of bakery products is relatively simple compared to the elaborate methods needed for many other foodstuffs, because either they are distributed by systems that take into consideration their short shelf life or they are low-moisture products that do not readily support microbiological activity. Consequently, breads and rolls are often packaged in nonrigid containers such as paper bags or (more commonly) heat-sealed plastic bags that offer moderate protection against environmental hazards such as insects, mold spores, and dust. Films used for the bags are often made of polyethylene or plastic-coated cellophane.

Samuel A. Matz

Cereal Products

Cereal grains are the primary ingredient of breakfast cereals. The processing of cereal grains into ready-to-eat or quick-cooking specialty items has been rapidly developed as a food industry, initiated only during the past century. These cereals have been promoted based on nutritional and health aspects, convenience, and variety. Consumer acceptance of breakfast cereals has been greatly influenced by product eating qualities, including size and shape, crispness, and flavor, all of which are controlled through the formulation and process sequences.

Early traditional prepared cereals were primarily precooked and dehydrated whole grains or kernel portions. Many basic cereals are prepared in near-traditional methods; however, most ready-to-eat products are prepared by blending flours (wheat, corn, oat, soy, and rice), malt, sugar, milk solids, and selected functional and nutritional additives with water, and mixing to produce a plastic dough that may be manipulated to yield distinctive products. Doughs are heated to produce a soft textured gelatinized starch and to provide desired color and flavor. Extrusion technology has been developed that permits continuous controlled blending and mixing, heating, and a broad spectrum of intricate size and shape configurations. The formed dough piece may be delivered at a critical moisture level or may require tempering prior to additional processing. Efficient extruders may be used as the complete heating and forming and expansion unit, or may be used as a specialized stage of preparation in a flaking or puffing operation. The continuous processing of a wide variety of cereal items can be accomplished only through critical control of formulation and process parameters.

Flaked cereals. The preparation of flaked cereals is initiated by relatively simple cooking of milled or whole grain.

Corn flake products. These flaked cereals are traditional. Blends of corn grits (free of germ and bran) are pressure-cooked in rotating horizontal cookers in the presence of water, malt, sugar, salt, and flavorings. This process requires an extensive cooking at about 18 lb/in.^{2} (120 kilopascals) steam pressure for 1–2 h to provide for starch hydrolysis and gelatinization, development of nonenzymatic browning, inactivation of enzymes, and infusion of moisture. Properly cooked grits will be tender and a translucent brown, possessing about 33% moisture. Cooked grits are mechanically separated and gently dried to about 20% moisture and tempered under conditions which will assure equilibration of moisture throughout each grit. Following the batch pressure cooking, a pelletizing process may be utilized to incorporate additional ingredients to enable uniform blending of a variety of grains into each flake, and to enhance uniformity and size of final toasted flakes. Tempered grits or pellets are passed between smooth cool flaking rolls. The gap (nip) between these rolls is controlled to aid flake size, thickness, and shape. Within the flaking operation, each grit is flattened and expressed to the characteristic flake shape. Tempered moisture control is important to assure uniform flake sizing. Preheating of tempered grits immediately preceding the flake rolls has been demonstrated to reduce cell rupture during flaking, resulting in a reduced rate of hydration (increased bowl crispness) in the final product. Cool moist flakes are hot air–toasted at temperatures

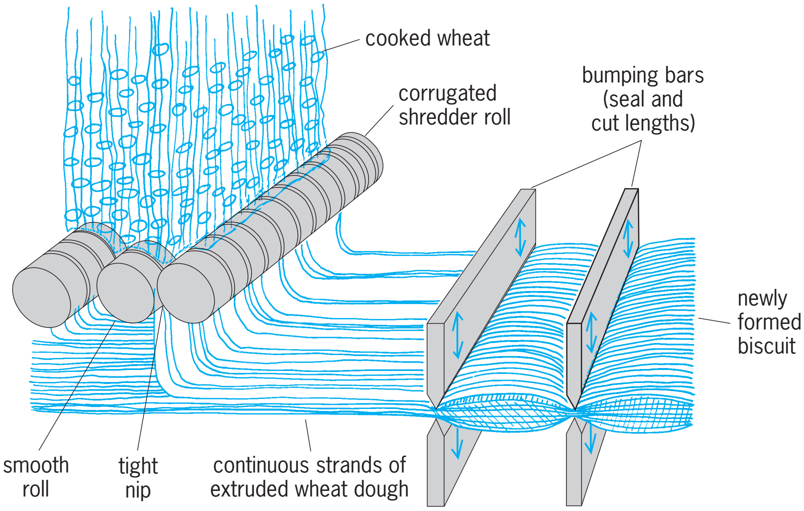


Fig. 8. Schematic for production of shredded wheat biscuits. (*After G. A. Leveille and M. A. Uebersax, Fundamentals of food science for the dietician: Wheat products, dietetic currents, vol. 7(1), pp. 1–8, Ross Laboratories, Columbus, Ohio, January–February 1980*)

exceeding 525°F (274°C) for differential times sufficient to dehydrate, crisp, brown, and blister the flake surface. The conditions of the toasting phase may be controlled to enhance flake crispness and surface blister formation. The final cooled toasted flakes may be spray-coated with an emulsion of vitamins and minerals, flavored or sugared, and dried to about 3% final moisture prior to packaging. *See* CORN.

Wheat flake products. These are prepared from whole, steam-tempered grain that is initially bumped through rollers to rupture the bran and increase moisture permeability through the endosperm. These kernels are pressure-cooked similarly to corn grits to about 50% moisture. Separated and partially dried kernels (30% moisture) undergo further air drying (20% moisture) prior to flaking. Dried cooked kernels are preheated and hot-flaked at 190°F (88°C) to yield a plasticized flake for toasting. Low-temperature (320°F or 160°C) zoned toasting enhances flake curling and surface blistering. *See* WHEAT.

Puffed cereals. The expansion and dehydration of cooked grain kernels or formed doughs is essential to impart the desirable textural properties to the final product. Control of expansion (decrease in bulk density) may be attributed to moisture, product formulation, and expansion process. Generally, expansion or puffing may be accomplished by gun puffing, or direct extrusion. Each requires controlled heating to facilitate rapid volatilization of internal moisture. Gun-puffed cereals, including puffed rice and wheat, are prepared by placing cooked grain (polished or pearled) into sealed vessels which are rapidly heated and pressurized up to 200 lb/in.^{2} (1400 kPa). Sudden release of pressure to the atmosphere expels and expands the cereal. Gun-puffed expanded cereals are also prepared from tempered dough–formed pieces. The expanded product will require additional drying or toasting. Oven puffing, commonly applied to cooked kernels or formed doughs, consists of continuous movement through a fluidized-bed heating chamber. High-velocity turbulent air lifts the cooked moist particles and rapidly heats them. The internal volatilized moisture expands the product volume, cools the surface by evaporation, and dehydrates the product. Critical control of oven design (staged heating zones), temperature, product characteristic formulation, and moisture are required to provide adequate expansion.

Extrusion systems can accommodate mixing, heating, and pressurizing ingredients in a continuous manner. The hot gelatinized dough held under high pressure will rapidly expand upon exiting the shaping die. This moist, expanded, continuously cut product requires further drying prior to packaging.

Shredded biscuits. Whole grain wheat is fully cooked and tempered to yield soft, pliable gelatinized kernels of approximately 50% moisture. This cooked wheat is formed into thin strands (0.04-in. or 1-mm diameter) by passing through high-pressure grooved rolls (**Fig. 8**). The continuous strands of pliable cooked wheat are stacked, generally up to 20 layers thick. The layered sheet is cut into biscuits by bumping bars set at designated intervals. Cut biscuits are dried and toasted (baked) to desired color and crispness in continuous-pass ovens.

Hot-served cereals. Numerous milled grains including wheat (farina), oats (regular rolled oats), and cornmeal (grits) are common. These products are steamed or toasted, dry-milled, sized, and packaged. Each requires cooking with water until palatable (generally more than 5 min). Development of rapid-cook or instant products has received much attention. Instant farina has been prepared by using selected proteolytic enzymes that will reduce cooking times to 1 min. Quick-cooking oats prepared from steel-cut thin-rolled groats are widely available and also cook within 1 min. Instant oatmeal is rendered palatable in the bowl by the addition of boiling water. This product is prepared by blending thin-rolled flakes and a hydroscopic gum to aid rehydration. Numerous fruits or flavored ingredients have been incorporated to aid in this product's popularity. *See* CEREAL.

Mark A. Uebersax

Milling cereal grains. Seeds from cereal crops and pulses provide a wide range of food products, including staples such as bread, pasta, couscous, and dhal. In their natural state, seeds are covered by a fibrous protective coating (seed coat, bran, or husk), which sometimes contains antinutritional substances that protect the seed against predators or disease-causing organisms. Removal of the outer layers of the seed (dehulling or decortication) often improves the nutritional or organoleptic quality of the food product. Dehulling of grain can be achieved by abrasion, attrition, impact, or roller milling, depending on the architecture of the seed. Dehulled grain is reduced to flour in a variety of grinding mills. Wet milling or air classification can further break seed components into starch- and protein-rich fractions, which have specialized uses in the food and chemical industries. *See* GRAIN CROPS; LEGUME.

Abrasion milling. Abrasion milling is used for grains such as sorghum, millet, or rice where the seed-coat layers adhere tightly to the endosperm. Typically, vertically or horizontally mounted, rotating grinding stones gradually rub off the outer layers of the grain, thereby exposing the seed endosperm. Small-scale abrasive dehulling machines (**Fig. 9**) have been gradually introduced in many developing countries to eliminate the time-consuming mortar-and-pestle process. After dehulling, the grain is removed from the machine by tipping over the dehuller barrel. Subsequently, the powdered seed-coat layers are separated from the dehulled grain by winnowing or sieving before grinding the latter to flour. Batch processing of grain in this type of machine ensures that a small-scale farmer in a developing country is still able to consume the grain variety that is produced on the family farm. *See* MILLET; RICE; SORGHUM.

Attrition and impact milling. These methods are used on types of grain that have thicker seed coats that are not tightly bound to the seed (for example, field peas and soybeans). An attrition mill removes hulls as the grain passes between two steel or stone disks rotating in a horizontal or vertical plane (similar to a plate mill). One of the disks may be stationary, and the surfaces of either disk may be configured in a variety of contours. In impact milling, the seeds are hurled at high velocity against an abrasive or metal surface, to crack the seed coat. *See* PEA.

Roller milling. Roller milling is commonly used on wheat to simultaneously separate the bran layers from the endosperm and to reduce the latter to flour. Roller mills are also used to produce cracked grain, as well as wheat and corn flakes. A roller mill consists of a pair of cast-iron rolls with a hard, wear-resistant surface. Adjusting the gap between the rolls produces different degrees of grinding. A wheat flour mill usually contains a series of differentially rotating break and reduction rolls. Break rolls, which are always corrugated, serve to initially cut and shear the wheat seeds into large pieces. Reduction rolls, which have smoother surfaces, further reduce the particle size and scrape the endosperm from the bran layers. Although the modern roller mill has not undergone any profound change, manufacturers have improved the drives, bearings, and feed mechanisms. The most important improvement in wheat flour milling has been in terms of automation to reduce the labor component of the cost of production. For example, in Japan a fully computerized mill with a capacity of about 440 tons (400 metric tons) of wheat per 24 h is operated by a single miller during each 8-h shift.

Combined techniques. An innovative method of combining abrasion- and roller-milling techniques has been developed as a means of improving the recovery of wheat flour in the milling process. Following tempering of the wheat with small amounts of water for a short time, about 65% of the bran is initially removed by friction and abrasion in modified rice polishers. The remaining bran, which is located in the crease of the wheat, is subsequently removed during conventional roller milling. It has been determined that semolina yields are approximately 7% higher when durum wheats are milled with this technology.

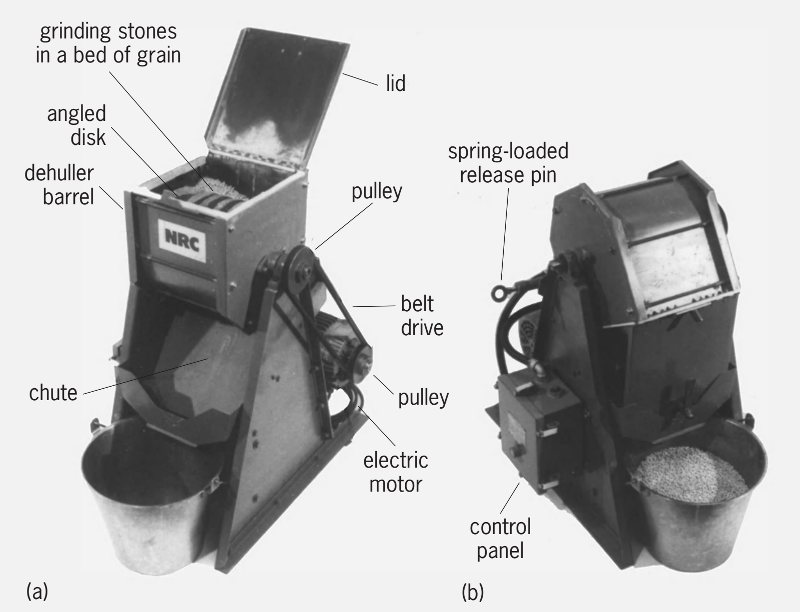


Fig. 9. Small-scale abrasive dehulling machine in (*a*) upright and (*b*) dumping positions. (*Plant Biotechnology Institute, National Research Council of Canada, Saskatoon*)

Development. Food engineers that develop and test milling equipment work closely with plant breeders to ensure that new grain varieties can be milled efficiently. Several small-scale mills have been developed which are able to accurately predict milling performance of gram quantities of early-generation plant breeding material. R. D. Reichert

Confections

It is difficult to define confections because the forms in which they appear are extremely diverse. One common factor is that they are consumed primarily for the sensory properties of appearance, flavor, and texture rather than for their nutritional value. Most of them contain large percentages of sweeteners, such as sugar, corn syrup, or honey. Generally, bakery products such as cakes and cookies are not considered to be confections, although they may have many of the same characteristics, and may be used as one of the components of candy bars (as when sugar wafer cookies are sandwiched with caramel and enrobed with chocolate).

Candies can be classified on the basis of their structure and texture as being primarily dependent on either sugar, protein, or fat for their predominant characteristics. All candies contain sugar or some other sweetener, but hard candies (glassy or amorphous forms of sugar with very low moisture content) and fondants (soft masses of very small sugar crystals surrounded by saturated syrup) rely on the characteristics of sugar for their basic structure. Proteins such as egg white, gelatin, casein, and soy protein concentrates form the matrix of caramels, marshmallows and many other types of foamed candies, and nougats. Starch-and-gum-based confectionery, with their typical soft chewy texture, form another category. Chocolate is the prototypical form of fat-based confectionery, relying on a melt-in-the-mouth reaction rather than solubility in saliva for the pleasant textural sensation that differentiates it from the sugar and protein structures. The final form of candy depends as much on the processing technique as it does on the formula. *See* COCOA POWDER AND CHOCOLATE; SUGAR.

Hard candies. These confections consist of up to 95% sweetener, with colors, flavors, and often other characterizing additives. They vary in form from the highly colored, almost transparent fruit ball to ribbons or sticks having very complex decorations and containing fillings of contrasting texture and appearance. The sweetener, which is usually a mixture of sucrose and corn syrup, is in a glassy or noncrystalline state, and the candy must contain a very low amount of water (typically, about 1%) to prevent the sugar from crystallizing during storage.

In the original method for preparing hard candy, a concentrated solution of the sweeteners is boiled in an open bowl heated by a gas flame until most of the water is evaporated and the temperature reaches about 310–330°F (154–166°C). Great care is taken to ensure that all crystals are eliminated during cooking; if any solid material accumulates on the sides of the kettle, it is washed or brushed into the boiling candy. Any crystals left in the molten sugar will promote crystallization throughout the entire mass as it cools, ruining the texture and appearance of the finished candy. When the correct temperature is reached, the molten mass is dumped onto a cooling table, and colors, flavors, and acids are mixed in. The temperature will be around 260°F (127°C) at the start of this operation, and in the 200–220°F (93–104°C) range at the finish. Different-colored rods of material can be assembled on the cooling slab to form a pattern that will appear in reduced form in the finished candy pieces.

While it is still plastic, the hot candy block, weighing perhaps 50 or 60 lb (23–27 kg), is gradually formed into a continuous thin rod by rolling and pulling. The batch roller into which the block is placed supports the large, irregularly oval shaped lump of cooling candy on either four or six conical rollers that rotate first in one direction, then in the other. The cavity containing the rollers is heated to keep the sugar plastic. From the batch roller, the thick rope of candy is drawn through the sizer, which consists essentially of a series of heated pairs of rollers having concave and knurled surfaces. Each set encloses a smaller opening and rotates at a faster speed than the preceding pair of rollers. From the sizer, the rope passes into formers, which are available in various designs.

There are several devices for large-scale batch processing and automatic cooking of hard candy formulas. Many of these devices subject candy that has been heated to around 280°F (138°C) to a vacuum of 16 in. mercury (54 kPa) or greater to remove most of the remaining water. Quality is improved, because the less stringent heat treatment minimizes sugar degradation to give improved clarity and keeping quality. Vacuum cookers are often designed to spread the hot syrup in a thin film on steam-heated surfaces to rapidly evaporate the water, which is swept away by hot air currents. The concentrated syrup is then exhausted into a receiving vacuum chamber where the final cooking is performed. In continuous systems, the candy is discharged from the vacuum chamber through a pump or air lock into a hard candy–depositing system or a cooling conveyor. When it is deposited onto a belt, the colors and flavors are metered onto the cooked sugars, the strip is rolled and folded by plows, and the final pieces are formed substantially as described above. For automatic depositing plants, additives are placed in the molten sugar before it is fed into the pump which deposits measured quantities into metal or plastic molds. After a brief cooling period, the molds expel the finished candy pieces onto a conveyor leading to the packing equipment, and return to the depositer for filling. Although they are very efficient and produce uniform pieces of excellent quality, the continuous systems lack versatility and flexibility.

Fillings, such as chocolate, jam, or peanut butter, can be injected by a variable-speed positive-displacement pump into the candy as it is being formed into a rope on the batch roller. As the diameter of the cylinder is gradually decreased through the successive sizing rollers, the proportion of filling to coating remains approximately constant. The squeezing together of the candy rope as the mold closes around it seals the softer filling inside a casing of hard candy. When hard candy pieces are being made by the deposit methods, fillings can be introduced by using concentric nozzles, each fed by a separate metering pump. Filling is extruded through the inner nozzle at the same time that the molten sugar casing is extruded from the external nozzle.

Gums and jellies. These confections have relatively high moisture contents and, as a result, they are elastic and soft and have comparatively short shelf lives. A common variety is the orange slice, which, like most of the types in this category, is basically a starch gel sweetened with sugar and corn syrup and having a moisture content of 18–20%. A typical formula is 33% granulated sugar, 49% corn syrup [64 D.E. (dextrose equivalent), an indication of the sweetness of the syrup], 3.3% thin boiling starch, 4.9% high-amylose starch, and 9.8% water. Although open-pan cooking was the customary procedure in earlier times, continuous processing is more rapid and convenient and is used for most production in the United States. In one such method, the ingredients are first mixed and preheated to about 200°F (93°C), then passed through a jet cooker held at 335°F (168°C) to complete starch gelatization. The cooked product is exhausted into a mixer where the flavor and color are added, or these ingredients are metered continuously into the jet cooker effluent.

Pieces are usually formed by depositing the paste into molds, but the oldest method involves spreading the jelly onto a cooling table and cutting out segments after the starch sets up. It is customary to "sand" the pieces, or cover them with coarse sugar, to reduce their tendency to stick together.

Hard gums, normally very firm, sticky candies, can be based on gum arabic as the gel agent and contain the usual sweeteners, colors, and flavors. Since starch gelatinization is not a requirement, the sweetener solution is simply boiled to around 225°F (107°C) to establish a satisfactory concentration, and then it is mixed with gum arabic that has been previously softened in water. The finished blend is deposited in starch molds and dried for several days until a satisfactory texture develops. The pieces may be either steamed or oiled to give them a glossy surface.

Pectin, agar, and gelatin can also be used with sweeteners and other ingredients to form confectionery jellies. Different textures are obtained. Gelatin tends to produce toughness, agar usually contributes a "short" texture, pectin gives soft tender jellies, while starch is more versatile. Modified starches have been developed specifically for use in confectionery. *See* AGAR; GELATIN; PECTIN; STARCH.

Products based on crystallized sugar. Fondant, a suspension of very small sugar crystals in a supersaturated sugar syrup, forms the basis of many confections such as cream centers for chocolates. Fondant can be prepared by dissolving sucrose along with some invert sugar or corn syrup (for example, 75% sugar and 25% corn syrup) in water and evaporating enough of the moisture to bring the water content to about 12%. At this point, the solution is supersaturated with respect to sugar, and caution must be taken to avoid premature crystallization. If the solution is beaten vigorously while it cools, the excess sugar will tend to be deposited in the form of minute crystals of relatively uniform size. It is generally agreed that crystals measuring more than about 35 micrometers in their largest dimension will be detected as a roughness or coarseness when the fondant is consumed—a negative organoleptic characteristic. Good fondants contain crystals falling mostly in the range 10–15 μm.

Fondant with added colors and flavors can be used as the basis for several types of confections, but it is normally too dense for optimum acceptability. A lighter, softer texture can be obtained by mixing the fondant with a frappé. This component is prepared by mixing a foam promoter, such as egg white, with sugar and corn syrup, and whipping the mixture to a meringuelike consistency. Typically, 7–10 lb (3–4.5 kg) of frappé is added to about 100 lb (45 kg) of fondant base. Whipping can be performed in the usual mixers found in bakeries, or in pressure whisks that enclose the mixing operation in a vessel containing compressed air. Continuous whipping machines are available.

Cream centers for chocolates can be shaped by extruding and cutting off in the desired piece size, but more elaborate shapes are formed by depositing the fondant mixture into cavities pressed into trays of cornstarch or into plastic molds. Liquid-center chocolates are often made by including a very small amount of the enzyme invertase in the formula. The fondant is liquefied during a period of days or weeks as the result of the hydrolysis of sucrose into the more soluble invert sugar by the enzyme. Since invert sugar is considerably more soluble than sucrose, the firm-cream center (which can be enrobed easily with chocolate) eventually changes to a thick syrup.

Butterscotch, nut brittles, toffees, and caramels depend for their characteristic textures and flavors on the interaction of sugar, dairy products, and fairly large additions of fat. Scotching is a term sometimes applied to the reactions that occur when sugar and butter are heated together until a light brown color develops. Butterscotch was probably the original confection formed in this way, and various types of toffees and caramels were developed by adding other ingredients and modifying the method. Flavor development in these products is due partly to a reaction between reducing sugars and proteins, and partly due to poorly understood reactions of the triglycerides.

Caramels can vary from soft (up to 10% moisture content) to hard (as low as 5% moisture content) and can have different degrees of graininess, which is affected by the size of sugar crystals. Fudges might be regarded as an extreme example of caramels, with high moisture content and a short texture resulting from a large amount of sugar crystallization.

Aerated products. Most of the types of candy described above can be expanded by air bubbles, forming new or modified types of confections. Some examples are nougat and marshmallow. Even hard candy, which is almost pure sugar, can be aerated by stretching and folding the hot mass [about 220–210°F (93–99°C)] through many cycles. Air is entrapped and subdivided into small bubbles by this procedure (which can be performed by hand or automatically) so as to give an opaque "satin" appearance and a more tender texture. Taffy pulling is another example of this type of processing. Heat-softened pieces of hard candy can also be expanded in vacuum chambers to give puffed balls of low density. The centers for malted milk balls are made in a similar fashion.

Nougats are made by whipping a mixture of high-boiled sugar syrup, a milk ingredient, fats, and a foaming agent such as egg albumin or modified whey. A soft, slightly chewy confection results, with the texture dependent to a large extent upon the amount of air which has been whipped into the confection. *See* ALBUMIN.

Marshmallows can vary from very soft to very firm in consistency, but they all consist of a high percentage of sweetener with some structure-forming material such as gelatin to support the foam that is established when the mixture is whipped. Most marshmallows are formed continuously by passing the mixture through pressurized mixing chambers in which assemblies of vanes rotate rapidly. Common cylindrical marshmallows are formed by extrusion and cutting, but the variety known as circus peanuts and similar versions are deposited into starch molds and allowed to partially dry before they are packaged.

Panning. Nuts and various kinds of small candy pieces can be coated with layers of sugar in a revolving pan. Chocolate can be treated in this manner to give multicolored crunchy-coated disks of candy. The old-fashioned jawbreakers were made in this manner, with many layers of different colors being applied to each piece. The pans are bowl-shaped with an opening in front, and range in size from small lab models to production equipment 4 ft (about 1 m) in diameter. Most of them are rotated by an axle that is inclined about 30° from the horizontal and is powered by a variable-speed drive. They can be heated by steam jackets or gas flames. Up to about 200 lb (90 kg) of candy pieces are poured into the opening as the pan rotates. Small amounts of adhesive solution and powdered sugar or other coating materials are poured or sprayed into the bowl at intervals. As the candy pieces tumble and spin inside the drum, they pick up the coating material in thin, uniform layers. When the coating has been completed, the pieces can be dried by hot air currents. The pieces can be coated with a final layer of carnauba wax or the like to give a shining finish. Separate pans can be used for the finishing or polishing operation, and they may have internal metal ribs to speed up the process. When coating with chocolate, the melted chocolate is sprayed into the pan in small amounts, with each layer being solidified by cool air currents before the next portion is added. Most chocolate-covered pieces are finished by coating with a polishing syrup composed of dextrin, sugar, and corn syrup or with a lacquer dissolved in alcohol. *See* CARNAUBA WAX.

Chewing gum. This is a special kind of confection which consists of a matrix of insoluble latex or gum (natural or synthetic) containing, as a discontinuous phase, sweeteners, softeners, flavors, colors, and other additives. As this mixture is chewed, microscopic pockets of flavoring are gradually exposed to saliva and dissolved. The original gum base was chicle, obtained from the sap of a tree native to Central America. Later, jelutong from Indonesia and sorva from South America were used as replacements for chicle. In modern manufacturing, large quantities of synthetic gums, resins, and waxes are used as part or all of the gum base ingredient in formulations. They often have superior textural properties and "staying power," and are more uniform than natural latexes. Among these synthetic ingredients are poly(vinyl acetate), synthetic rubber of the polyisoprene type, polyethylene waxes, and resin esters. *See* CHICLE.

Production of chewing gum starts with the blending of chunks of gum base (often presoaked with some of the liquid ingredients) and the other formula components in a steam-jacketed heavy-duty is 19.4% gum base, 19.8% corn syrup, 59.7% powdered sugar, 0.5% glycerine, and 0.6% flavor. The mixture will usually be discharged at 113–131°F (45–55°C) and formed into loaves weighing 8–10 lb (3.6–4.5 kg). After cooling, the gum is passed through pairs of rollers in series to gradually reduce its thickness. Powdered sugar may be applied to the surface to prevent sticking. The final set of rollers scores the strip of gum into stick size, and cuts it into sheets which are placed on trays and conditioned at 59–64°F (15–18°C) and 45–50% relative humidity for 24–48 h. In the packaging line, the sheets are broken into individual sticks, wrapped, and packaged. In large factories, the process is continuous up to the conditioning stage, which may take place in a continuous tunnel or in controlled-atmosphere compartments. Bubble gum, which has a different base composition, does not require conditioning and is sized into ropes which are immediately cut into balls or other pieces. Samuel A. Matz

Condiments

Condiments are aromatic or savory vegetable substances which contribute distinct taste or character to foods, change or enhance specific flavors of foods, and modify the appearance of foods by adding color, as with paprika, saffron, or turmeric. The individual nature of a condiment depends on the type and concentration of volatile oil and other chemical constituents.

Condiments include spices, herbs, flavoring agents, chemicals, sauces, and other materials. One way of classifying them is according to the part of the plant from which they originate: (1) roots such as garlic, ginger, sassafras, horseradish, and turmeric; (2) stems and leaves, such as basil, sage, tarragon, thyme, bay leaf, and marjoram; (3) barks such as cinnamon and cassia; (4) buds such as capers, cloves, and saffron; (5) fruits (ripe or immature) such as paprika, pepper, capsicum, juniper, and vanilla bean; (6) seeds such as mustard, cardamon, anise, celery, dill, coriander, bitter almonds, and poppy. Other condiments may be classified as chemicals (monosodium glutamate, sodium chloride, sodium citrate) and sauces (wine concentrates, vinegars, relishes).

Properties of most condiments are distinct because of their volatile characteristics or the substances of definite chemical composition therein.

With the increasing interest in international cuisine, or ethnic foods, the use of condiments has increased tremendously and should continue to do so. Many of these cuisines accent savory seasonings, spices, and penetrating tastes and aromas, such as in Italian, Cajun-Creole, Mexican, and Asian foods. With the use of savory spices which complement basic taste and flavor, the amount of salt can be reduced, a positive medical attribute. Commercial spices are frequently sold as blends. They are often irradiated prior to packaging, and frequently are marketed in concentrated form.

Standards have been established by the Federal Food and Drug Administration for the more important spices, herbs, and seeds as listed above. These consist of specifications for percentages of crude fiber, total ash, volatile oils, starches, bacteriological limits, and other basic constituents of each specific item. The purpose of such standards is for the maintenance of uniformity of product, the insurance of freshness, and the provision of better quality control for the consumer. *See* MONOSODIUM GLUTAMATE; SALT (FOOD); SPICE AND FLAVORING.

Desserts

Desserts are generally considered to be sweet foods that are served at the end of a meal. Commercial preparation of desserts includes the following methods: canning, freeze-drying, drying, production of dried powders, and aseptic techniques. In spite of the emphasis on dieting and weight reduction, processed desserts are found in greater variety and amounts than ever. However, single portions are available to assist in dieting regimes.

Canning is used mainly for fruits. The fruit is packed in sugar syrup or in sugarless water solution—with or without synthetic sweeteners in order to reduce calories, or for dietetic purposes. The use of saccharin is being replaced by use of aspartame and acesulfame-K, which are white, odorless, crystalline products 200 times as sweet as sugar. Thickeners, such as vegetable gums, are added to simulate sugar syrup. Fruits of excellent visual appearance are often packed in glass. The canning process is also used for puddings, such as rice pudding or Indian pudding; and crepes suzettes and baba au rhum are canned by using syrups or semithick bases.

Frozen desserts include fruits such as strawberries, rhubarb, peaches, and mixed fruits, including exotic tropical fruits. Specialty items such as pies, cakes, puddings, crepes, and many types of ice cream, sherberts, and related desserts are also frozen. *See* ICE CREAM.

The development of a combination of freezing and drying, that is, freeze-drying, has produced superior fruits that are very close in taste and texture to the natural fruit.

Dried fruits, such as figs, apricots, plums, and grapes, are traditional desserts that were prepared by sundrying. This method is coupled with improved drying methods that provide controlled time, temperature, and humidity.

Gelatin desserts are made by combining sugar, plain gelatin, citric acid, sodium citrate, flavor, and color. Pudding powders, such as chocolate pudding, may be regular (requiring cooking) or instant (requiring mixing only). These contain powdered sugar (dextrose), starches (pregelatinized in instant pudding), flavor, and color. Puddings have been developed that are made with products such as tofu (bean curd) or yogurt, as well as flans that contain carrageenans (Irish moss). Additional developments include high-fiber muffins and cakes that are healthful desserts and alternatives to eating breakfast cereals. *See* AGAR; CARRAGEENAN.

By using aseptic techniques, puddings, such as rice pudding, tapioca, or chocolate, are packed in individual containers, either plastic or cans, and then they may be stored at refrigerated or room temperatures.

Another method for packaging baked goods has been developed using inert gases; this method provides flavor retention and increases shelf life. It is known as controlled-atmosphere packaging (CAP) or modified-atmosphere packaging (MAP). These baked products are usually packed in microwavable polystyrene trays. Alfred J. Finberg

Soft Drinks

Soft drinks are nonalcoholic beverages, that is, they contain no so-called hard liquor such as whiskey or gin. Soft drinks can be divided into two classes: carbonated soft drinks (sometimes referred to as soda, soda water, and soda pop) and still beverages (also referred to as noncarbonated soft drinks, fruit-flavored drinks, and fruit punches).

Soft drinks were first offered for sale during the seventeenth century in Europe as a mixture of water, lemon juice, and honey. Their origins can be traced to the attempts by many Europeans to copy and improve the carbonated waters and beverages that were known for their therapeutic value. The first company of *Limonadiers* in Paris sold their lemonade from tanks which were carried on their backs. The production of soft drinks has evolved from this stage to highly sophisticated production lines that are capable of filling up to 2000 cans per minute. The packaging has also evolved from returnable glass bottles to lightweight nonreturnable glass bottles, and to steel, aluminum, and plastic containers. In many communities the glass bottles and aluminum cans are being segregated from other solid waste and recycled. There are studies under way involving use of plastics that are recyclable or biodegradable after being discarded.

Carbonation is the main difference between the two classes of soft drinks. Carbonated soft drinks are prepared by dissolving carbon dioxide in water which is to become part of the soft drink, or in the completely formulated soft drink itself. A volume of carbon dioxide gas will be absorbed by an equal volume of liquid at 60°F (15°C) and 14.5 lb/in.^{2} gage pressure (100 kPa); correspondingly, at 60°F (15°C) four volumes of carbon dioxide gas will be absorbed by the liquid at 45 lb/in.^{2} gage pressure (310 kPa), and five volumes at 60 lb/in.^{2} gage pressure (413 kPa).

Except for the carbonation, there is very little difference between the different classes of soft drinks. All normally contain a sweetener, acidulant, preservative (if necessary), and flavoring. The nutritive sweeteners (sugars) used may be sucrose (table sugar) derived from sugarcane or sugarbeets, 50% inverted type consisting of 42–47% sucrose with the remaining percentage equally divided with fructose and glucose, or with high-fructose corn syrup (HFCS). Regular corn syrup has not been found to be suitable for soft drinks; however, 42% high-fructose corn syrup (approximately 42% fructose and 58% glucose) and 55% high-fructose corn syrup (approximately 55% fructose and 45% glucose) are acceptable. The exact amount of dry sugar solids may vary in a product depending on the type or combination of sweeteners that is used. *See* FRUCTOSE; GLUCOSE.

The combination of sugars available to the soft drink manufacturer allows flexibility should there be a disruption in the supply of one sweetener type or if economic factors dictate the decision to purchase one over the other. There are also advantages and economy of time in the production process when using liquid sweeteners versus granulated cane or beet sugar, which must be dissolved first.

The acidulant is normally citric, malic, or phosphoric acid, depending on the product. Citrus-type beverages (for example, orange) usually use citric or malic acid. The colas as a rule use phosphoric acid. When necessary, the preservatives are either sodium benzoate or potassium sorbate, depending on the pH of the beverage. Not all beverages contain preservatives. Soft drinks are products very rigorous to degradation. They usually avoid microbiological spoilage because of their anaerobic environment and low pH. Preservatives are normally used only as an added precaution or when the product formulation warrants it. The flavoring is usually one of the ingredients added in the smallest amounts. Flavor chemistry has advanced to the stage where the flavor component is highly concentrated, and this provides savings in the costs of transportation and packaging.

Optional ingredients may also be added in such proportions as are required to accomplish their intended effects. These could include fruit juice, antioxidants, foaming agents, and so forth. Diet beverages are sweetened with saccharine alone $2.4–4.2 \times 10^{-4}$ oz/fl oz (0.24–0.41 mg/cm^3) or in combination with aspartame. Many products are sweetened with 100% aspartame alone. Both saccharin and aspartame are nonnutritive sweeteners. All ingredients must be food-grade or higher to ensure the highest quality and a uniform product. A trend has developed in the soft drink industry to use as many natural ingredients as possible in the product formulation.

Although potable water is generally taken from a municipal supply, it is processed further to eliminate or reduce impurities that could affect the quality of the finished product. In a few soft drink manufacturing plants, the water treatment equipment may simply consist of a sand filter and an activated carbon purifier. The sand filter is sized to allow a flow rate of approximately 2 gal/ft^2 (81 liters/m^2) per minute. This filtration removes minute matter. The carbon purifier removes not only chlorine and color but also tastes and odors that may be present. In most soft drink plants the water is first treated chemically by using an oxidant (such as chlorine or chlorine dioxide), lime, and ferrous sulfate. The lime serves two purposes: it reduces the water alkalinity by the removal of the calcium and magnesium bicarbonates, if necessary, and it serves to convert the ferrous sulfate to a ferric hydroxide floc. The flocculent masses entrap very minute particles, including algae and planktons. The water is then passed through the sand filter and carbon purifier. The carbon purifier is really another filter that contains activated carbon which removes color, taste, and chlorine. *See* WATER TREATMENT.

Mark J. Pietka

Sauces

The word sauce is derived from the Latin *salsa*, meaning salty or salted, and is defined as a condiment or combination of condiments in liquid or semiliquid state, eaten with food as a garnish or accompaniment for the purpose of adding flavor, piquancy, or interest. Sauces may vary in thickness from light to heavy. Some are thin liquids, with finely dispersed colloidal particles, while others may contain large pieces of solid materials.

Preparation. The use of sauces developed from the ancient method of food preservation by salting. Sauces function in three basic ways: to enhance flavor, to "atone" for its absence, or to act as a binder for other ingredients. Every sauce has a binding ingredient to hold the liquid portion and provide required texture. There are two principles in preparing sauces: the roux principle (a mixture of flour and butter) and the emulsion principle (egg yolks used to absorb and suspend particles). There are three distinct types of roux, from which mother sauces are created: white, blonde, and brown, and these give rise to hundreds of variations. The white roux is a thoroughly blended mixture of flour and butter. Blonde roux is a white roux cooked until it turns a straw color. Further cooking until nut brown produces a brown roux. Oil or another fat can replace the butter in a brown roux.

Three basic sauces prepared from blonde roux are béchamel sauce (white sauce), velouté sauce, and sauce parisienne (sauce allemand). Two basic sauces are derived from brown roux: brown sauce (sauce espagnole) and, with the addition of a meat jelly, demiglace.

There are three basic sauces based on the emulsion principle: hollandaise sauce, mayonnaise, and vinaigrette sauce, also known as French dressing.

Commercial. Many commercial sauces are available; some are processed and are stable at room temperature until used by the consumer. Others are refrigerated or frozen at the time of manufacture and must be stored at proper temperatures. The frozen products are usually more expensive but frequently are judged to be superior to products processed in jars or cans.

Packaging. The most common packaging materials are glass and plastic, since these make the product visible. Tubes, similar to toothpaste tubes, are common in Europe, but have not had wide consumer acceptance in the United States. Unit portions in individual foil or plastic packets are a convenient method of dispensing and serving sauces and dressings in fast-food chains and on planes. Advances in laminated foil packaging have increased shelf life.

Standards. The Food and Drug Administration has established standards of identity for mayonnaise, French dressing, and salad dressing. These standards state allowable ingredients and their minimum amounts. There are no official standards for other sauces and dressings except that all ingredients must be listed on the label in order of amount.

Specially formulated dressings that are low-calorie, low-fat, or low-sugar have been developed as a result of consumer interest in preventive medicine, nutrition, and weight control. The result has been an upsurge in interest in reduced-calorie mayonnaise and dressings. To qualify for low-calorie status, the caloric content must be reduced by at least one-third, one serving supplying not more than 40 calories (167 joules).

Manufacturing. Commercially prepared sauces are either liquid (water-based) or emulsified (oil-based). The aqueous sauces, such as hot sauce, mint sauce, and Worcestershire sauce, may be stabilized to minimize separation.

Improved methods of processing have been incorporated into the handling, cooking, homogenizing, and packing of manufactured sauces. Tomato sauce and tomato ketchup are made from strained tomato pulp concentrated with vinegar, sugar, and spices to different solids levels.

In ketchup manufacture, pulped tomatoes are concentrated by boiling or in a vacuum concentration unit to the desired consistency and solids content. The condiment is made into several grades, depending upon the concentration of the tomato solids. Fancy ketchup has more than 33% of total solids present, while standard grade has as little as 25%. The prepared condiment is preserved by canning, usually in glass containers for consumer uses. Chili sauce is tomato pulp with the seeds remaining and concentrated with added sugar, vinegar, and spices to specific solids level.

Worcestershire sauce is a combination of vinegar, hydrolyzed proteins, vegetables, and fruits that is allowed to age in barrels to develop a balanced flavor blend. Soy sauces contain proteins such as wheat, corn, or soybean processed by controlled acid hydrolysis, with condiments, molasses, and caramel color added. Pepper or hot sauce is made by concentrating chili or tabasco peppers and adding vinegar and salt. Meat or barbecue sauce is a combination of tomato pulp, soy base, vinegar, and spices. Mustard sauce and prepared mustards are made by grinding or milling mustard seed and adding vinegar and spices; novelties include various blends, for example, “sweet and spicy” or “sweet and hot.” Spaghetti sauces, combining tomatoes, herbs, and spices, with or without meat, are processed in cans or jars. Following the increase in popularity of ethnic foods, particularly Italian pastas, many variations of tomato sauce have been introduced.

Applesauce is a combination of cored, sliced apples cooked with sugar, water, and spices. Cranberry sauce is made by cooking cranberries, sugar, and water to specific solids level. There are many types of fruit sauces, such as those made from peaches and apricots, which are concentrated, combined with sugar and spices, and packed for use as ham glazes or with meat dishes. Wine is a component in some special sauces.

Sauces that are generally emulsified are mayonnaise, salad dressing, and related dressings. They are oil-in-water emulsions with a vegetable oil content ranging from 30 to 85%. The oils used are mainly soybean or cottonseed. These products also contain 4–10% egg yolk, with the exception of the pourable types of French dressings, which generally contain a vegetable gum or stabilizer to maintain emulsification. They are commercially manufactured with mills of colloid or homogenizing types, pressure types of homogenizers, ultrasonic homogenizers, or continuous automatic mixers combined with high-pressure units. An improvement in this field was the introduction of instantaneous cooking followed by continuous cooling of the starch phases of salad dressings. The average commercial salad dressing contains 30–40% vegetable oil. The difference is made up with cooled starch paste for salad dressings. The jet-type handling of starches has effected tremendous economies and streamlining of the entire salad dressing industry.

Quality. Digital processing and computer controls have provided further efficiency and quality control, leading to longer shelf life and product stability. The coordination of these developments has resulted in uniform final emulsification by proper dispersion due to particle-size reduction control. Commercial products such as hollandaise and bearnaise sauces have been direct results of these applications used with sterilization procedures.

Soups

A soup is generally a thin puree or liquid consisting of an extract of meat or other ingredients (such as vegetables) plus seasoning. It is prepared by boiling in a single pot. Soups normally are classified as basic consommes or bouillons, thin broths, or thicker soups. Modifications are made by thickening with tapioca, flour, cream, egg yolks, or other solids such as pureed vegetables, legumes, pieces of fish, poultry, or meat.

Originally, simple purees of tomato were the most popular commercial soups available to the public. Sophisticated consumer tastes and manufacturing skill have made soups like gazpacho, she-crab, and cream of cheddar cheese commonplace, and the variety is increasing constantly. Soups satisfy the modern consumer market needs, because they are easy to prepare and readily digested. They are economical and suitable for all ages, health conditions, and types of weight control.

Canned. Early commercial soups, typically tomato, chicken, vegetable, and meat soups, were heat-processed in cans. Usually these were in concentrated form (50% concentrate), with water, milk, or a mixture of both to be added in home preparation. Some soups were canned single-strength and served without dilution. Today many variations, such as ethnic soups, special-diet soups (for example, low in salt), and single-portion sizes are available in cans.

Because of convenience to the consumer, newer types of soups that can be opened, heated, and served in the same container within a few minutes have become available; however, the largest number of soup products available are the canned variety.

Dry soup mixes. These are packed in flexible bags made of foil or paper boxes which are automatically sealed. The mixes are usually composed of dehydrated soup stock and may include dehydrated vegetables, meat, chicken, and spices mixed together sufficiently to obtain uniform distribution. Onion, chicken-noodle, and beef soups are typical of the varieties available.

Instant soups. These may be made from freeze-dried ingredients, spices, and precooked noodles, which rehydrate upon addition of boiling water. Unlike dry soup mixes, these require no cooking and are very useful for rapid preparation.

Frozen soups. These soups are cooked with a shorter process time than that for canned soups. After cooking, they are packaged and then frozen. Varieties best suited for freezing are fish-and-vegetable combinations such as chowders and barley-mushroom, where flavor and texture of the ingredients are better retained by short process time.

Microwavable soups. Because more people work and have less time for food preparation, this development is beneficial. These soups can go from the shelf or freezer, to the microwave oven, and to the table in less than 10 min. Some varieties are packed in sturdy high-density polypropylene bowls. Most of these soups are in frozen form, although some are shelf-stable and packed in plastic.

Compressed cubes. These were an early commercial type of soup that are still popular. They are made of salt, meat or chicken, and vegetable extracts with spices, and are slowly soluble in hot water to provide a clear bouillon. Alfred J. Finberg

Fermentation

In the broad sense, fermentation is defined as a metabolic process in which organic compounds undergo chemical changes as a result of enzymatic actions. The participating enzymes, specialized types of proteins that function as biological catalysts in chemical reactions, are synthesized by all living organisms, including bacteria, yeasts, and molds. Since a broad spectrum of plant and animal materials readily supports the growth of fermentative microorganisms, fermentations are universal mechanisms used to prepare and preserve numerous foods and food products. *See* FERMENTATION.

Alcohol fermentation. Fermenting yeasts share a common function in the manufacture of beer, ale, and wine, namely, the production of alcohol. Hops (dried blossoms of the hop plant and an essential flavor ingredient in beer) is added to the water extracts of germinated barley grains, boiled, cooled, and inoculated with selective strains of top or bottom types of fermenting yeasts, thereby producing beer or ale, respectively. In wine production, the sugars of grape juice are readily available for direct fermentation by yeasts. Following fermentation, the liquor is removed from the sediment, aged, and filtered to achieve ultraclarity. Red and white table wines are produced from singular grape varietals or blends of varietals and usually contain 10–13% ethyl alcohol, whereas dessert wines are prepared by fortification, that is, they contain added amounts of alcohol and sugar. *See* ETHYL ALCOHOL; MALT BEVERAGE; WINE; YEAST.

Dairy fermentations. Cheeses and cultured dairy drinks (yogurts, buttermilk, and so on) are produced by the acidic fermentations of milk by lactic acid bacteria. Various types of cheeses are prepared from curds (coagula formed by acidification) that have been heated at 122°F (50°C) for varying periods of time. Soft-type cheeses (cottage, cream, Neufchatel, and so on) are prepared from curds that have been subjected to reduced cooking times, whereas increased holding times followed by extended periods of ripening at 35–60°F (2–15°C) produce the cured cheeses such as Cheddar, Parmesan, and provolone. Cofermentations, that is, processes in which mixtures of lactic and nonlactic acid bacteria or molds produce cheeses which possess unique flavors and distinct physical characteristics such as the eye or hole formation in Swiss cheese (gas-producing bacteria) or the blue-vein effects found in bleu or Roquefort cheeses (mold). Process cheeses are prepared by heating blends of cheeses and emulsifiers to a smooth mass which is subsequently dispensed in containers or packaged as sliced cheeses.

Vegetable fermentations. Fruits and vegetables (for example, cabbage, carrots, cucumbers, or olives), when treated with salt or brines, readily support lactic acid fermentations. In sauerkraut fermentation, granular salt is thoroughly mixed with shredded cabbage, 2–3 lb (0.9–1.4 kg) of salt per 100 lb (45.4 kg) of cabbage, whereas with intact fruits and vegetables, such as cucumbers or carrots, the commodities are totally immersed in the brines prior to fermentation. The levels of lactic acid generated during fermentation are governed by the amounts of fermentable sugar present in the plant materials; fermented pickles usually provide less than 1% acid, whereas more than 2% acidity is produced in the fermentation of sauerkraut. During processing, fermented sauerkraut and brines are heated to 170°F (77°C) prior to container closure, whereas pickles, upon removal from the fermentation tanks, are thoroughly washed to remove excessive amounts of salt, packed in containers containing formulated brines (spices, sugars, vinegar, and so on) and pasteurized at 165°F (73°C). Pasteurization can be circumvented by incorporating chemical preservatives into the packing brines. The latter products may be packaged in glass containers or flexible plastic bags and stored at refrigerator temperatures.

Pickling. Brines, containing vinegar (notably acetic acid), spices, and condiments, are commonly used for preserving and imparting unique and distinct flavors to many foods. Many of the fermentable foods described above can also be prepared by nonfermentative pickling processes. For example, fresh cucumbers immersed in brines overnight are packed into containers and covered with hot brines containing 2–3% salt, vinegar (1.5% acetic acid), and spices. The acidified product is then heated at 165°F (73°C) for 15 min, followed by prompt cooling. In addition to plant materials, fish (such as herring or salmon) and cured meats (such as bacon, ham, or sausage) can also be prepared by the pickling process. John R. Stamer

Fish and Seafood Products

Fish and shellfish landed by American fishing boats are used for both industrial products and food. The industrial products are fishmeal and fish solubles used primarily as poultry feed and fish oil. The food products are fresh and frozen fish, canned fish, and cured fish (salted, smoked, and pickled). The following descriptions of processing methods emphasize American methods. *See* MARINE FISHERIES.

Industrial products. In the United States a greater volume of fish is used to make fish meal and fish solubles for animal (primarily poultry) feed and for fish oil than goes into any other human food operation. This is because the United States menhaden catch, none of which is used for human food, often is greater than that of all other finfish combined. *See* ANIMAL FEEDS.

Menhaden are rendered by the wet process (**Fig. 10**). The catch is unloaded by pumping with water, and after draining and weighing, fish pass into the raw box to await processing. The fish are cooked in a long, horizontal, cylindrical steam-jacketed cooker; often these have openings to inject steam into the cooker, which contains a screw conveyor to move the fish through. Emerging from the cooker is a mixture of solids and liquid consisting of oil and of water in which nitrogenous and other substances are dissolved. These materials go into a press and emerge as a press cake that passes into steam-jacketed dryers to produce the meal to which an antioxidant, usually butylated hydroxytoluene (BHT), is added. This retards oxidation of the residual oil to prevent overheating, which otherwise might start a fire during meal storage. The liquid phase from the press goes through centrifugal separators to give oil and stickwater, the latter being a solution of protein and other solubles. The stickwater is condensed in multiple-stage evaporators, reducing moisture content to 50% or less. The oil is purified in oil polishing centrifuges.

In addition to the very large production of meal and oil from menhaden, meal, oil, and solubles are by-products derived from tuna cannery waste. Processing procedures are very similar to the procedure used for menhaden.

Some whole fish, especially those of low oil content, are rendered into meal by the dry process; that is, the fish are dried directly without cooking or pressing.

Fresh and frozen fish. Fish to be marketed fresh or frozen usually are sold in one of the following forms (**Fig. 11**): whole, drawn, dressed, headed and gutted, steaks, or fillets. In small-scale operations, the whole or dressed fish may merely be washed, sometimes scaled, packed in boxes of crushed ice, and shipped to wholesale or retail markets.

In large-scale operations, especially if fillets are to be cut, semimechanical packing lines may be employed. The fish are first washed and sometimes scaled in cylindrical open metal washers provided with sprays of water. The fish then pass along moving belts to work stations where they are cut into fillets. The fillets next pass through a washing tank, sometimes followed by a dip in tripolyphosphate solution, especially if the fillets are to be frozen. Fillets may then pass over a translucent glass plate, lighted from below, to be inspected for the presence of parasites or defects. The fillets then go to a packing table for hand packing and sealing in moisture-vapor–proof containers. Containers are filled carefully in order to avoid air pockets, which would result in oxidation, discoloration, and off-flavor development.

Freezing of fillets often takes place on shelves made of pipes that contain the refrigerant. Sometimes fans circulate the air. In other cases the packaged fish are frozen between plates containing the refrigerant, with considerable pressure on plates and fish (plate freezing); or the packages are placed in air tunnels with rapidly blown refrigerated air.

When whole or dressed fish are frozen and stored, they are protected against dehydration and oxidation by dipping after freezing in cold water to form a protective ice glaze. If the fish are packed in paper-lined wooden boxes, they may require only infrequent reglazing; otherwise, they are frequently reglazed by spraying with water.

Fish steaks, a form used especially for salmon, halibut, or swordfish, are cut with a band saw after freezing and ordinarily are glazed before packaging.

Breaded fish products. The production of breaded fish portions is an important manufacturing operation. Fish fillets are frozen into blocks, ice-glazed, and shipped to points where the breaded product is to be prepared. The blocks are cut on band saws into either stick form or nearly square portions. The product then moves along a belt where first a batter is applied, and then they are covered with a dry breader. The final color of the product depends upon the type of batter and breader used. Commonly batters contain grained corn flour, corn meal, and nonfat dry milk solids, while breaders are often winter wheat cereals, bread, or cracker crumbs. The products are then cooked for several minutes in hydrogenated oil or liquid corn oil at 375–405°F (191–207°C). An alternate product is breaded but not cooked. The cooked products must be cooled slowly and packed while they are at 90–100°F (32–38°C).

Canned fish. Tuna and salmon are the two most common American canned fish (close to 90% of the total fish pack). Formerly tuna was packed largely in southern California. This has changed, and nearly all of the tuna is packed by American companies at canneries closer to the locations where the tuna are caught, primarily in American Samoa and Thailand in the Pacific and Puerto Rico in the Atlantic. Salmon are canned largely in Alaska, with much smaller amounts canned in Washington and Oregon.

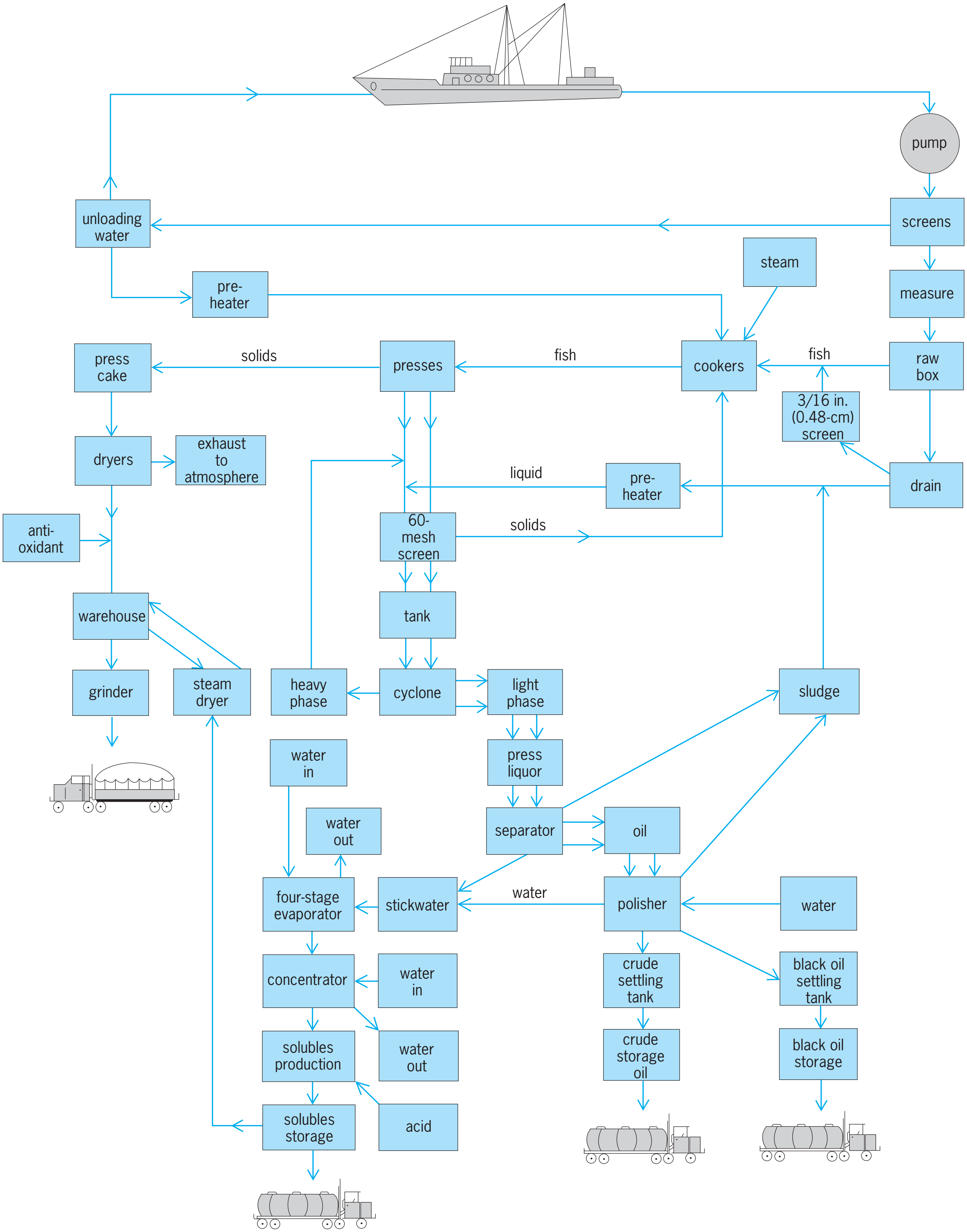


Fig. 10. **Flow chart for the preparation of fish meal, condensed fish solubles, and fish oil from whole fish by using the wet process. (*Zapata Haynie Corp.*)**

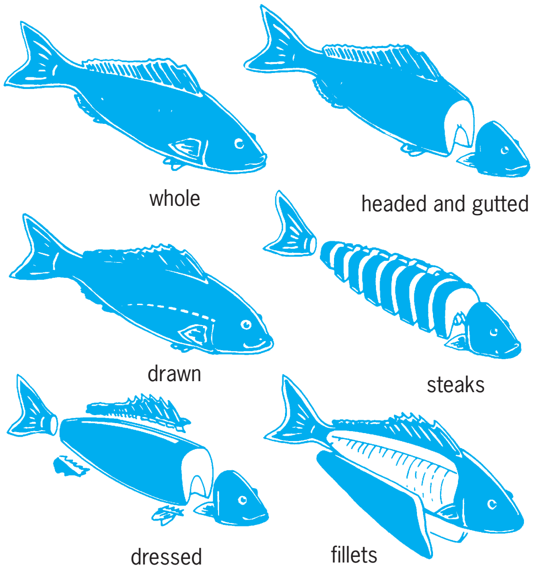


Fig. 11. Common market forms of fresh or frozen fish.

In the tuna-canning process (**Fig. 12**), the fish, most of which have been frozen on the fishing vessel and held frozen in air until approaching the tuna cannery, are completely thawed after unloading. They are then eviscerated by hand and inspected for removal of any fish showing indication of spoilage. The fish are loaded according to size in paper-covered wire baskets and precooked in ovens with steam at atmospheric pressure. The length of the precooking process varies with the size of the fish. Fish from the precooker are cooled, often overnight, and then cleaned by hand to remove skin and bones and to separate the dark from the light meat. Dark meat and other trimmings are used for pet food and for making meal and oil. The tuna loins are cut and packed by machine as either solid meat tuna or chunk pack. The filled cans pass down the line where salt and then oil—generally soya bean oil—heated to about 180°F (82°C) are added. The cans are then sealed and retorted.

Salmon are usually delivered unfrozen to the cannery, where they are sorted as to species. In the cannery the head is removed and the fish slit to remove eggs, which are exported to Japan. Dressing, which includes removal of viscera and cleaning, is done by means of a machine known as the iron chink. The fish are scaled and washed at the sliming table and then enter a cutter that chops fish into can-sized chunks that pass into the filler machine, which inserts them into cans. Salt is added by machine at levels of 1.25–1.67% salt. Cans then enter a machine which rejects those that are underfilled. At the patching table, underfilled cans have additional salmon added. Filled cans with lids loosely attached then enter the vacuum sealer. After passing under sprays of water, the cans are retorted, then cooled.

Cured fish. Unlike the situation in many other countries, curing (salting, smoking, pickling) is no longer a major fish preservation method in the United States. Salting of fish is mainly used for preparing mild cured salmon, an intermediate step toward smoking. Smoked fish (important species are salmon, sablefish, chub, herring, whitefish, and whiting) as well as small quantities of a large variety of pickled herring products are produced primarily for areas such as greater New York, Chicago, and Los Angeles. There has, however, been a steady decline in demand for all these cured fish products.

Innovations in processing. Fish flesh can be obtained in maximum yield by using special equipment such as the type used for many years by the fishing industry in Japan, where minced fish is in demand (**Fig. 13**). Such cheaply obtained minced fish can be used for products familiar to the American consuming public, such as fish fillets molded into shape from minced fish and used in the production of frozen fish blocks for the preparation of breaded fish products. Also, minced fish can be molded into shape and sold as artificial crab legs. Such uses of minced

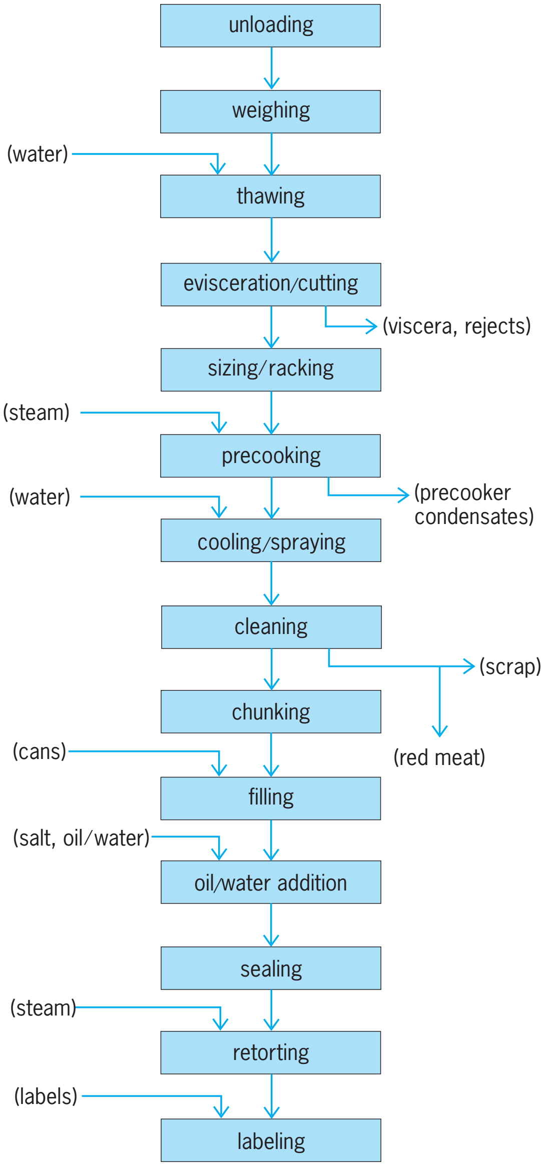


Fig. 12. Processing steps in the canning of tuna.

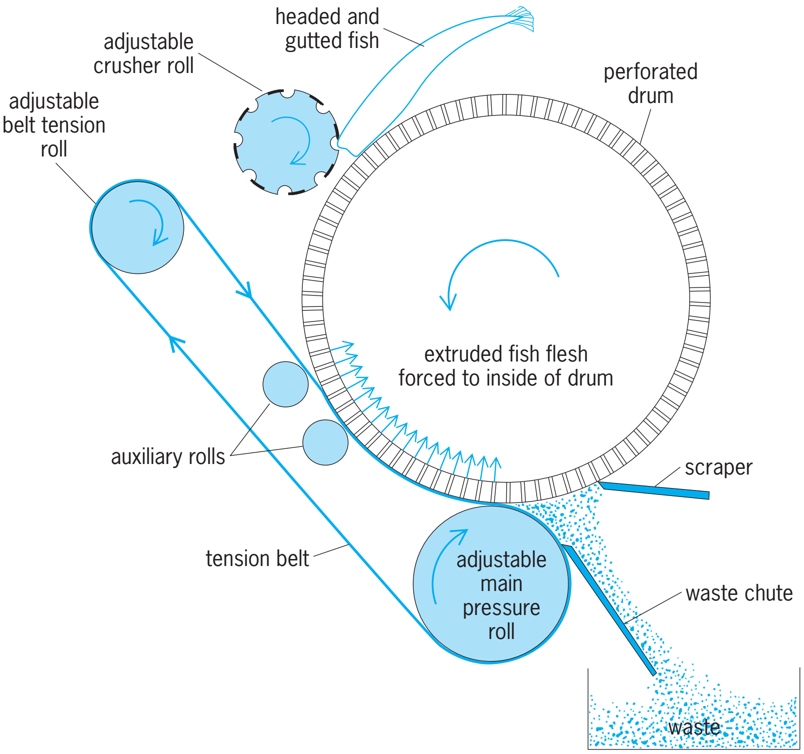


Fig. 13. Schematic sectional view of a fish flesh-separating machine to produce minced fish.

fish yields savings in production costs and, in addition, permits marketing of species unfamiliar to the public.

Another trend has developed from awareness by consumers of certain desirable nutritional properties of fish. Many fish are frozen in brine or refrigerated in chilled seawater, either of which allows penetration of varying amounts of salt into the fish. Federal regulations could require a statement of sodium content on package labels and that it be fairly uniform in amount among different samples of the particular product. The processes being used for fish preparation result in wide variations in sodium content, well beyond what would be permitted; some modification in processing methods is required.

Another nutritional property of fish relates to the findings that omega-3 fatty acids in the oil of fish seem to minimize greatly the occurrence of heart attacks. These desirable effects are due to the presence of certain highly unstable fatty acids that occur only in fish. In order for the public to gain maximum benefits from this property of fish, improved methods of packaging or treatment are being developed to ensure optimum retention of these beneficial substances. Maurice E. Stansby

Fruit Products

Botanically, a fruit is the ripened ovary and accessory parts of a flower. In common usage of the term, a fruit implies a fleshy product typical of trees or bushes that is high in acidity and sugar and, when ripe, possesses a characteristic flavor; however, many vegetables (such as tomatoes, squash, and beans) are true fruits. Fruits, such as apples, peaches, pears, and plums, are consumed primarily for their fine texture and distinctive flavor, as well as their high nutritive levels of vitamins and minerals. Wide differentiations in flavor, color, and texture exist among fruit varieties, some of which are best suited for fresh market distribution while others are appropriate for processing. *See* FRUIT.

Handling fresh fruits. Fruits are highly metabolic and thus normally undergo relatively rapid postharvest degradation in terms of quality. Respiration, a natural biological process (see reaction below), results in consumption of sugar and oxygen and in production of carbon dioxide and heat, but will subsequently result in fruit senescence.

$$\underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Oxygen}}{6O_2} \xrightarrow{\text{enzymes}} \underset{\text{Carbon dioxide}}{6CO_2} + \underset{\text{Water}}{6H_2O} + \underset{\text{Heat}}{673\ \text{kcal}}$$

Ethylene gas, which accelerates the respiration and ripening rate of many fruits, may be used under controlled conditions to enhance the storage life of that fresh fruit. Respiration rate is also dramatically influenced by temperature; thus, low temperature (41°F or 5°C) may extend high-quality shelf life. Several fruits, however, are cold-sensitive and undergo chilling injury, which reduces their quality. This moisture loss is decreased by maintaining high relative humidity (90–95%) conditions. Wilting and shriveling of fruit, caused by the process of transpiration, are a result of water loss during storage. *See* ETHYLENE.

Controlled-atmospheric storage. Significant progress in long-term storage of fruit under specified and monitored conditions has been achieved. Apples are particularly well suited for controlled-atmospheric storage. Fruit is sealed in warehouses equipped to maintain controlled temperatures, relative humidity, and gas (atmosphere) environments. Maximum storage life is attained by lowering oxygen concentrations to approximately 2% and increasing carbon dioxide (CO_2) levels to about 5%. Scrubbing the air to remove ethylene gas also improves storage conditions. High-quality shelf life, as evidenced by firm, crisp, and flavorful fruit, may be increased by many months with these controlled-atmospheric conditions.

Canning. Unlike vegetables, which require pressure-cooker processing (250°F or 121°C) to inactivate spores of *Clostridium botulinum,* fruits naturally possess sufficient acidity (low pH) to permit thermal preservation at boiling water temperature (212°F or 100°C). Fruits are prepared by washing, sorting, and trimming. Specialized preparation may be required for some fruits, such as stem and pit removal for cherries and peel and core removal for apples. Sliced or prepared fruit may require blanching to remove tissue gases prior to filling into enamel-lined cans or glass jars. Aggregate fruit products (whole berries or sliced portions) may be canned in water or sugar syrups. Filled containers are exhausted by thermal or vacuum to remove air prior to hermetic sealing to assure removal of oxygen and adequate vacuum within the container upon cooling. Sealed containers are heated in water or atmospheric steam to achieve temperatures of about 200°F (93°C) in the container center prior to cooling. High quality of processed fruit can be achieved only through rapid handling, controlled heating, and proper selection of varieties.

Apple sauce. Peeled, cored, or whole chopped apples are heated with steam to inhibit oxidative browning and to aid in extraction of the fleshy pulp. The cooked mass may be sweetened and then passed through a fine-mesh sieve. This process, termed finishing, is typically accomplished by continuous extraction through a cylindrical screen in which rotating internal paddles express the cooked apples through 0.06-in. (1.5-mm) perforations. Large seeds and other firm portions that do not pass through the sieve exit through the end of the finisher.

Freezing. Freezing is effective in maintaining flavor and color qualities of many fruits. Rapid freezing techniques reduce ice crystal damage and improve quality. Properly prepared fruit may be frozen in sugar syrups or dry sugar packs (1 part sugar to 3–5 parts fruit) to reduce oxidation. Addition of ascorbic acid (0.2%) provides further control of pigment oxidation. Individually quick-frozen products are prepared by fluidized-bed freezing in which the fruit moves over high-velocity cold air currents to achieve rapid crystallization. They are then packaged and stored at −0.4 to −8°F (−18 to −22°C) as are all frozen fruits. Fruit purees, sugared or concentrated, are slush-frozen in a scraped surface heat exchanger, packaged as crystallized slurry, and then solid-frozen for later use in ice cream or baked goods.

Juice. Fruit juice products provide economic utilization of blemished but sound fruit. Fruits are initially washed and sorted and then prepared to facilitate maximum extraction. Fruit may be chopped and directly extracted (cold break) or heated to 180°F (82°C) prior to extraction (hot break). The latter method will increase extraction of pectic substances, inactivate pectinases (enzymes which degrade pectin), and produce a thicker juice possessing more body. Juice extraction is typically accomplished with batch hydraulic presses or continuous-screw presses which express juice and produce a dry pulp residue, termed pomace. Juice may be cloudy, that is, containing suspended solids, and may be further clarified by use of either pectinases or fining agents, such as bentonite, and centrifuged and filtered to remove the suspended material, producing a product of sparkling clarity. Juices may be pasteurized by continuous-flow heat exchangers to achieve 176–194°F (80–90°C) temperatures for 25–30 min (high-temperature–short-time). Pasteurized juice is suitable for canning in tin or glass, freezing, or refrigerated distribution. Sodium benzoate added at 0.1% is also effective in stabilizing acidic fruit juices.

Grape juice, produced by either hot or cold pressing, requires precipitation and removal of crude tartrates prior to bottling. Chilling or slow freezing will accelerate crystallization and sedimentation of tartrates. Following these pretreatments, the detartrated juice is suitable for packaging.

Citrus juices are produced by press or screw extractors, deoiled by vacuum distillation, and centrifuged, finished, deaerated, pasteurized by the high-temperature–short-time method, filled at high temperature and sealed, and rapidly cooled. Citrus concentrates utilize vacuum concentration to remove water without imparting dramatic deleterious flavor changes; however, add-back processes are used to dilute the evaporated concentrate with sufficient fresh juice to give proper final solids content and flavor profiles. Citrus concentrates are frozen in consumer packages requiring three parts water for reconstitution or are used for preparation of reconstituted, single-strength, refrigerated juice for local distribution. Advances in aseptic packaging technology have shown great potential for production of high-quality fruit juice products for use in institutional and retail markets.

Dehydration. Sundrying is traditional and accounts for a large volume of commercially dehydrated fruit. Prepared fruit is dehydrated following treatment to control the enzymes that cause oxidative browning. Blanching (steam), treatment with sulfur dioxide (sodium bisulfite solution), or drying in the presence of burning sulfur controls browning and enhances color and flavor of the dried product. Hot-air–drying is most common with fruit placed in trays for cabinet dryers (batch) or tunnel dryers (continuous) in which product moves on carts, trays, or mesh belts. Moisture is removed to about 20% to ensure microbial stability. Fruit "leathers" are prepared by drum-drying purees, while instant fruit juice powders are produced by spray-drying systems. Dehydrated fruits are packaged to control moisture absorption during storage. *See* DRYING.

Jams, jellies, and preserves. Production of pectin-acid-sugar gels from fruits is a traditional means of improving utilization of blemished but sound fruit. Jellies are prepared from extracted juices; jams from whole purees; and preserves from whole, intact fruit. Pectin (either naturally present or added) in the presence of sufficient sugar (increased hydrogen bonding), water, and acid (pH 2.9–3.4) forms a delicate gel. Standard jellies are prepared using 45 parts of fruit juice and 55 parts of sugar and concentrated to 65% soluble solids (65° Brix scale, a hydrometer scale for sugar solutions), and are hot-filled and cooled in final package (typically glass or plastic). Control of formulation and process is essential to attain a properly set gel. Dietetic jellies are produced using pectin containing limited methoxylation (pectic acid) which, in the presence of calcium, cross-links to form a gel without requiring sugar. Special care in heating and deaerating fruit for preserves and the use of a rapid-set pectin are essential to maintain distribution of fruit in the gel, thereby eliminating the floating fruit defect. *See* GEL.

Vegetable Products

Vegetables are plant products obtained as leaves, roots, stems, flowers, or fruits. The pH of vegetables ranges from 4.5 to 7.0. Handling of fresh vegetables requires controlled low temperature and high humidity to retard desiccation. Vegetables may be classified according to their respiratory activity; high-respiring vegetables, such as peas, beans, and corn, have less postharvest stability than do low-respiring vegetables, such as potatoes, beets, or carrots. Selection of optimum harvest time (maturation stage) is essential if maximum quality is to be attained. This is frequently in opposition to maximum yield; therefore, quality grading procedures often have a function in pricing structures for harvested vegetable crops. Postharvest cooling of vegetables aids in the removal of field and respiration heat, thus improving quality. Hydrocooling in chilled water, forced-air cooling, or evaporative (vacuum) cooling are often used.

Canning. Vegetables are washed with chlorinated water to remove soil, stones, and other debris and to reduce the microbial population. Products are sorted, trimmed, peeled, or sliced according to specific needs. Peeling can be done by mechanical or chemical means. Carrots and potatoes may be peeled by mechanical abrasion or by chemical lye peeling, in which products are dipped into hot lye (5–10% sodium hydroxide at 180°F or 82°C) for a short time to dissolve the outer peel. Peels and residual lye are readily removed by water sprays.

Blanching is a relatively mild heat treatment to which a product is subjected before it is packed into the container. Blanching serves as a final wash to reduce microbial load, to remove certain raw or off-flavors, and to increase the pliability of the product, thus facilitating filling and permitting the expulsion of dissolved or occluded gases from the plant tissue. During blanching, the product is exposed to hot water or steam for several minutes. In hot-water blanching systems, the product must be submerged in water at about 194°F (90°C) for a set time. Direct steam blanching reduces nutrient losses and water requirements and has thus found wide use in the industry.

Products are put into containers immediately after being blanched. Filling systems range from simple, manual operations to highly engineered machinery. Products packed in salt brine or syrups are first put into containers and then covered with the liquid portion. Filled containers are generally exhausted by being passed through a hot water bath or steam chamber to expel gas and to expand the product before the can is sealed.

Hermetically sealed low-acid foods (vegetables) require processing temperatures in excess of 212°F (100°C) to inactivate *Clostridium botulinum* spores. These temperatures are commonly attained in pressurized steam retorts. Specific handling preparation procedures and process schedules are required for each individual vegetable product.

Commercial vegetable canneries frequently require transportation of crops from large geographic areas to enable an economically efficient operation by extending the normal local crop season. Many vegetable canneries also rely on dry beans, peas, or other dry, stable crops as an off-season production item. Dry beans require soaking (overnight in cold water or for a short time in hot water) to hydrate the seed and reduce splitting prior to filling.

Filled, exhausted, and sealed containers require pressurized cooking at 250°F (121°C) for approximately 40 min to assure health safety and produce a tender, palatable product.

Pickling. Lactic fermentation of vegetables produces a delicate flavor in a wide variety of products. Cucumbers are salt-brined and fermented by naturally occurring lactic bacteria. Salt content in the brine is measured in terms of the Salometer scale, 100° Salometer = 100% salt saturation. Salt concentrations are initiated in 1000-bu (3.5-m^3) wooden vats at 40° Salometer (around 10% salt). This concentration is reduced to about half this level in 1–2 days because of osmotic removal of water from the cucumbers. As the fermentation continues, salt concentrations are increased to 40–50° Salometer over several weeks.

Fermented salt stock is preserved by the acid and salt, and it may be stored for months prior to processing. Processed pickles require desalting by water extraction prior to other treatments, such as sugar impregnation to form sweet pickled products.

Genuine dill pickles are produced by natural lactic fermentations in the presence of spices and relatively low salt concentrations. Selected other vegetables (onions, peppers, and cauliflower) may be brined and preserved in similar fashion.

Large numbers of cucumbers are produced as fresh-packed products. Whole, cross-cut or longitudinally sliced cucumbers are filled into jars, flavored, and acidified with vinegar before steam pasteurization.

Sauerkraut is produced as a lactic fermentation of finally shredded cabbage. Salt (2.5%) is added to clean sound shredded cabbage, thoroughly mixed, and packed into fermentation vats. The distinctive lactic flavor is developed during several weeks of active fermentation in the brine produced from the cabbage tissue. Kraut is heated in its own juice, packed into glass or tin containers, and processed in boiling water.

Freezing. Block packing or individually quick freezing are suitable methods for preserving high-quality vegetables. Blanching pretreatments (176°F or 80°C water or atmospheric steam) are essential to inactivate enzymes. Improperly blanched products will undergo dramatic flavor changes during frozen storage, and the result is a grassy or an ensilage taste and odor. Rapid-freezing systems reduce ice crystal damage and enhance textural characteristics of the vegetable product.

Dehydration. Prepared vegetables (sliced, diced, or chopped) are suitable for hot-air–drying and are often used as ingredients for formulated and prepared products, such as soups, casseroles, or entrees. Cabinet, tunnel, or kiln dryers are commonly used to reduce moisture content to less than 4%. Vegetables require blanching prior to air-drying.
Mark A. Uebersax

Meat Products

The processing of meat begins with slaughtering and dressing of cattle, sheep, and hogs to produce meat for human consumption. The carcasses are normally chilled to near freezing temperatures to retard the growth of pathogenic and meat spoilage bacteria. Most carcasses are reduced into primal or subprimal cuts and distributed to supermarkets or other retail outlets, or else they are used for further processing. Beef primals and subprimals are generally vacuum-packaged in gas- and moisture-impermeable bags, packed into corrugated paper boxes, and shipped as boxed beef. This retards the growth of undesirable microorganisms and extends the shelf life of fresh meat. Although limited numbers of carcasses are distributed as halves or quarters with all cutting being done at the retail store, the amount of cutting at the retail level has steadily decreased. There has been an increase, however, in centralized cutting and packaging at warehouses that service some retail markets.

Further processing of meat involves one or more of the following procedures: freezing; curing; smoking or cooking; canning; sausage manufacturing; drying; production of sectioned and formed meat products; manufacturing of restructured meat products; enzyme tenderization; and rendering of fat. These processes decrease perishability and provide a variety of different products with distinctive flavor and texture characteristics that add to their desirability.

Frozen meats. Freezing greatly extends the storage life of meat. It is utilized not only for fresh raw meat but also for cooked meats and for many convenience items, such as precooked meat entrees, meat pies, and meat items in frozen dinners.

Although frozen meats are stable for months, their acceptability declines because of development of oxidative rancidity. In cooked products, rancidity develops much faster and is called warmed-over flavor. Development of rancidity can be delayed greatly by covering precooked meats with a sauce or gravy having antioxidant properties. Extracts from many vegetables possess natural antioxidants and can be used in the sauce or gravy.

Cured meats. Curing also greatly extends the storage life of meat products. It is accomplished by adding salt either alone or in combination with sodium or potassium nitrite. Sugar, polyphosphates, and spices may also be added to impart special properties to cured meats. The curing ingredients are commonly made into a brine and injected under pressure.

Salt acts by dehydration and altering osmosis to inhibit growth of spoilage bacteria. Nitrite retards development of rancidity and produces the characteristic cured meat flavor. It also reacts with the meat pigments to stabilize the color, and most importantly inhibits the growth of pathogenic bacteria, such as *Clostridium botulinum*, and other spoilage microorganisms. The level of nitrite allowed in curing is limited, since excess amounts may react with the natural amines of the meat and produce carcinogenic *N*-nitrosamines. Polyphosphates increase the water-binding capacity in cured meats. Sugar and spices add flavor.

Cured meats also may be smoked or cooked. However, some products (for example, corned beef and canned hams) are produced without smoking.

Smoked or cooked meats. Smoking meat is carried out for development of a smoky flavor, preservation, creation of new products, development of color, and protection from oxidation. Smoking and cooking are frequently carried out together.

Although smoke was originally generated by smoldering wood or sawdust, liquid smoke is now frequently used and this has led to a sharper demarcation between smoking and cooking. Liquid smoke is prepared by destructive distillation of wood. The particle phase is removed by washing with water, leaving the vapor phase in solution. Liquid smoke is applied by spraying, dipping, or atomization with heat. Its advantages are in uniformity of application, in ease of cleaning, and in preventing pollution.

Most smoked meats are cured and also cooked—including frankfurters and bologna, which can be eaten without further heating. However, fresh pork sausage may be smoked but requires cooking before being served.

Canned meats. Manufacturing of canned meats is accomplished by heat processing to produce either pasteurized or shelf-stable products. The former must be held under refrigeration, as is the case for most canned hams and picnics (bottom part of pork shoulders). Shelf-stable canned meats are heated for sufficient time at 250°F (121°C) to destroy most pathogens and can be stored without refrigeration.

Although more ham is canned than any other meat item, corned beef hash, beef stew, chili con carne, Vienna sausages, meat balls with gravy, potted meat, and canned luncheon meat are all common commercial products. Some canned meat products are reacted with curing ingredients before being sealed in cans and heat-processed.

Sausages. Meat products that are comminuted or emulsified with salt and seasoning are called sausages. Generally they are stuffed into cylindrical artificial or animal casings. Sausages may be coarsely or finely chopped or passed through an emulsifier to produce an emulsion, in which the particles are coated with either myosin or collagen proteins. On heating, the collagen gelatinizes to form jelly pockets, whereas myosin coagulates on cooling to form a stable emulsion. Since myosin is a contractile protein coming from skeletal muscle, an adequate quantity of lean meat is essential. Overchopping can also result in fatting-out as a result of too much surface area from the fat. *See* COLLAGEN; MUSCLE.

The most common sausage products include frankfurters and bologna, which are emulsion-type sausages; and these differ only in size. Meat loaves are also emulsion sausages but often contain added flavoring ingredients. Fresh pork sausage and smoked pork sausage are examples of uncured ground sausages, while thuringer and cervelat are semidry sausages and salami and pepperoni are dry sausages.

Dried meats. These products are produced by curing followed by drying at a low humidity. Drying must be carefully controlled, or surface case hardening will occur and retain the moisture in the center. These products include dried beef (jerky) and a number of dry sausages, such as pepperoni and salami. Their low moisture content makes them shelf-stable at room temperature.

Sectioned and formed meats. These products are produced from whole muscles or large chunks of meat that are bonded together to form a single fused piece resembling an intact cut. Binding is accomplished by (1) the myofibrillar proteins, which are extracted by tumbling or massaging to form a tacky exudate; (2) a meat emulsion, which is mixed with the intact pieces of muscle to fill the voids between pieces and to produce a sticky coating; and (3) adding a nonmeat binder, such as albumin, to coat the muscle pieces and bind them together. Salt is usually added to solubilize the binder. The meat pieces are forced into a casing to bring them into close contact. Heating then coagulates the binding material so that a single fused piece is formed during cooling.

Turkey rolls, sectioned and formed hams, and beef rolls are examples of these products. Usually sectioned and formed meat cuts are boneless, but some may contain a small amount of bone, such as mock chicken drumsticks.

Restructured meats. These products require specialized processing equipment. Three basic procedures are used: chunking and forming, flaking and forming, and shredding and forming. In chunking and forming, the meat is passed through a grinder or dicing machine to produce cubes no larger than 1.5 in.^3 (24.6 cm^3). Salt, polyphosphates, and other seasonings are added to the meat, which are mixed or tumbled to extract the myofibrillar proteins. Conversely, a small amount of a meat emulsion may be added as a binder. The chunked meat is then stuffed into casings or formed into meat logs that are frozen, tempered, and sliced.

Flaking is accomplished by high-speed equipment that cuts the meat into thin flakes. Salt is generally added to extract the myofibrillar proteins, which serve as binders. The meat is placed in a special press that forms meat logs which are frozen, tempered, and cleaved into steaks, chops, or roasts. These products are generally frozen and sold as frozen cuts.

Shredding and forming produces a texture more like intact meat cuts. The manufacturing procedures are otherwise similar to those used in production of restructured meat.

Restructured meats often suffer from oxidative rancidity. Limiting the time in frozen storage and decreasing salt levels can help in retarding rancidity.

Enzyme-tenderized meats. Fresh meats, especially the lower grades of beef, are frequently treated with enzymes of either plant or fungal origin to improve tenderness. The cuts are either dipped or injected with enzyme solutions and are generally frozen during distribution. Upon thawing, the enzymes act upon the myofibrillar and connective tissues and enhance tenderness. This process has made steaks of acceptable tenderness available to more consumers at a lower price. A patented enzyme tenderization process involves injection of papain into the bloodstream of cattle. The enzyme is then carried to the tissues by the circulatory system before the animal is slaughtered. The enzyme fraction is active only during cooking, resulting in an increase in tenderness.

Meat fats. These are valuable by-products of the meat industry and are produced by rendering (heating), which separates the fat from the supporting connective tissues. In manufacturing lard from pork fat and tallow from beef fat, decolorization, deodorization, plasticization, interesterification, fractional crystallization, and hydrogenation may be used to alter the characteristic appearance, color, odor, flavor, and firmness.

Rendered fat is classified as edible or inedible based upon its suitability for human food. Inedible grades are used for animal feeds or for industrial purposes, such as manufacture of lubricants or soaps. *See* FAT AND OIL (FOOD); SOAP. A. M. Pearson

Poultry Products

Poultry products in the United States include chicken, turkey, duck, goose, and game birds. The processed carcasses, or parts thereof, are supplied refrigerated as ice-packed or dry-packed, or as frozen products. The nonfrozen products are often called fresh, as opposed to frozen.

Commercial production and processing of poultry is confined to specialized production units that can deliver truckload lots of live birds (8000–10,000 broiler chickens) to processing plants. This results in uniformity in birds being processed at one time. Large plants for broiler chickens convert over 300,000 live chickens to chicken products in an 8-h day.

The conversion of live birds to poultry products involves a number of processes, all of which are controlled by a U.S. Department of Agriculture veterinarian inspector to assure consumers of the safety, wholesomeness, and compliance with label claims for all products. The birds are humanely dispatched and allowed to bleed thoroughly prior to scalding to facilitate complete removal of all feathers. The inedible viscera is removed along with the head and feet. The carcasses are thoroughly washed prior to chilling to less than 41°F (5°C) in an ice water bath. Some water is absorbed during washing and chilling, but this is carefully monitored by the inspection service personnel. Many of the processes are automated so as to keep cost to a minimum.

After chilling, broilers may be packed in boxes of about 24 carcasses with ice for refrigerated distribution to retail markets and institutional feeding establishments. Dry packing is in large plastic bags with about 24 carcasses per bag. The air is evacuated and replaced with carbon dioxide gas. The bags are sealed and then placed in boxes and distributed as refrigerated food at temperatures of less than 37°F (3°C). Distribution of poultry at 28°F (−2°C) has gained rapidly as a major temperature for handling. This deep-chill temperature keeps all free water in a package in solid form, thus greatly extending the shelf life. Some birds, especially turkeys, ducks, and geese, are individually packaged and frozen for distribution. Frozen poultry should be handled at less than 0°F (−18°C).

A large quantity of poultry products is cut up at the processing plant and packaged as complete birds (all cuts from one carcass) or as parts such as breast, thigh, drumsticks, wings, backs, and necks. Labeling requirements are very explicit on cut-up parts so as to accurately describe the portions of the carcass included in the cut, such as breast with ribs, or thigh with back. These requirements aid the consumer in getting exactly the portions desired.

Breast and thigh meat are often further processed by removing the skin and bone. The boneless breast meat may then be processed further into specialty gourmet items. The thigh meat is frequently cured and smoked to produce turkey or chicken ham. The sale of poultry meat as formed products (nuggets, patties, and croquettes) is an interesting development. Boneless breast, thigh, and leg meat are converted into the formed products. The uncleaned bones and some of the bony parts remaining are then mechanically deboned to yield mechanically separated meat. In many instances, some mechanically separated meat is used in formulations for the formed products. However, the mechanically separated meat is used as a major ingredient in the production of poultry frankfurters, bologna, or other processed poultry meat products.

As channels have developed for distribution of nonfrozen uncooked meat at 28°F (−2°C), a number of poultry processors are producing precooked poultry meat products packaged in microwavable containers for distribution at this deep-chill temperature. These precooked nonfrozen items are generally available from processors located near major metropolitan areas.

Egg Products

The proteins of the whole egg contain all essential amino acids in a balance favorable for meeting human dietary needs. An excellent source of all vitamins, with the exception of vitamin C, eggs also provide significant quantities of iron and trace minerals to the diet (**Table 3**).

Most eggs to be converted to products are broken by mechanical breakers capable of operating at speeds up to 100 cases, or 36,000 eggs, per hour. The machines can either separate whites and yolks or produce whole eggs at these production rates. The egg products (albumen, yolk, whole egg, or blends), standardized to specific values (**Table 4**), are then pasteurized prior to packaging for liquid or frozen distribution or for drying. Pasteurization lowers the total microbiological population and eliminates all pathogenic microorganisms. Times and temperatures of pasteurization vary with the products, the methods employed, and the requirements in different countries. A process known as ultrapasteurization has been developed for liquid whole eggs. It uses conventional pasteurizing equipment operated at higher than usual temperatures or longer than usual times to produce a pasteurized product in an almost sterile condition. By combining this with aseptic packaging, the shelf life of the liquid egg at 41°F (5°C) has been extended from days to weeks. It is anticipated that this process will allow movement of liquid whole eggs into a much larger market area.

Following pasteurization, eggs for freezing are packaged in 30-lb (14-kg) tin containers or in

TABLE 3. Nutrient composition of egg products, in grams per 100 g of eggs or egg products*

	Liquid or frozen				Dehydrated		
Nutrients and units	Whole egg	Egg white	Egg yolk, pure	Egg yolk, commercial	Plain whole	Stabilized whites	Plain yolk
Proximate							
Solids	24.5	12.1	51.8	44.0	96.8	93.6	97.2
Calories	152	50	377	313	600	388	692
Protein	12.0	10.2	16.1	14.9	47.4	79.1	39.2
Lipids	10.9	—	34.1	27.5	43.1	—	60.8
Ash	1.0	0.68	1.69	1.49	4.0	5.3	3.3
Lipids: fatty acids							
Saturated	3.67	—	11.42	9.16	14.51	—	20.35
Monounsaturated	4.60	—	14.67	11.80	18.18	—	25.64
Polyunsaturated	1.32	—	4.20	3.37	5.22	—	7.45
Cholesterol	0.37	—	1.18	0.96	1.48	—	2.11
Lecithin	2.32	—	7.20	5.81	9.16	—	12.84
Cephalin	0.46	—	1.42	1.15	1.82	—	2.54

*Except "Calories" category. 1 g = 0.035 oz.

SOURCE: Adapted from American Egg Board, *A Scientist Speaks About Egg Products* (1981), with the exception of cholesterol values, which are calculated from 1988 USDA values based on 213 mg per large egg.

TABLE 4. Specification guide for liquid, frozen, and dried egg solids*

	Liquid or frozen			Dried solids				
Category	Whites	Plain yolk	Whole	Angel whites	Flake albumen	Free- flowing stabilized whole	Free-flowing stabilized yolk	Fortified whole egg (tex) with carbohydrates
Moisture (max.), %	88.2	57	74	8	14	3	3	4
Fat, %	0.02 max.	25 min.	11 min.	None	None	40 min.	57 min.	31 min.
pH	—	—	—	7.0 ± 0.5	5.5 ± 0.5	7.5 ± 0.5	6.5 ± 0.5	7.5 ± 0.5
Protein (min.), %	10.5	16.0	12.5	80	75	44	30	27
Carbohydrate	—	—	—	None	None	None	None	30 max.
Color	—	SOP	SOP	Creamy	Clear yellow	NEPA 2–4	NEPA 2–4	—
Viable bacteria/g (max.)	50,000	10,000	10,000	5000	10,000	10,000	10,000	10,000
Yeast (max.)	10	10	10	10	10	10	10	10
Mold (max.)	—	10	10	10	10	10	10	10
Coliform/g (max.)	1000	10	10	10	10	10	10	10
Salmonella	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.
Granulation, %	—	—	—	100 USBS–80	—	100 USBS–16	100 USBS–16	100 USBS–16
Additives	—	—	—	SLS 0.1% max	—	—	SS 2% max	—
Performance	SOP	SOP	SOP	SOP	SOP	SOP	SOP	SOP

*SOP = specified on purchase; neg. = negative; SS = sodium silicoaluminate; SLS = sodium lauryl sulfate; NEPA = National Egg Products Association light-to-dark scale (range 2–4); and USBS–80 and USBS–16 = U.S. Bureau of Standards screens for determining particle size. 1 g = 0.035 oz.

4–10-lb (2–4.5-kg) wax-coated cardboard containers and blast-frozen at −20 to −40°F (−29 to −40°C). Egg liquid to be distributed as liquid is cooled to less than 40°F (5°C) and packaged in containers ranging from bulk tanks to the small cardboard containers used for frozen eggs. Egg liquid for drying is cooled to less than 40°F (5°C) and either dried on the premises or transported in insulated tank trucks to an egg-drying plant. Table 4 lists the properties of commercial liquid, frozen, and dried egg products.

Egg drying. Both liquid egg white and egg yolk from the breaking plants are collected at the drying plants. Prior to drying, the egg liquids undergo extensive processing to ensure retention of the functional and organoleptic properties in the finished product.

Egg whites. All egg whites for drying must be pasteurized. The small amount of free glucose present in egg white must be removed to prevent the Maillard reaction, which causes brown discoloration and off odors in the dried product. Glucose is removed by the use of bacteria or yeast capable of utilizing glucose or by enzyme systems capable of converting the glucose to a nonreactive material, such as gluconic acid. *See* MAILLARD REACTION.

Several whipping aids (most commonly sodium lauryl sulfate) have been approved by the U.S. Food and Drug Administration for addition to egg white in order to retain its high whipping qualities following drying. Following appropriate adjustment of pH with lactic, citric, or other approved food acids, the liquid egg white is spray-dried in conventional dryers. It is essential that egg white dryers be provided with dust-collecting systems. Spray-dried egg white solids are normally packed in 150-lb (68-kg) fiber drums or 25–50-lb (11–23-kg) polyethylene-lined boxes for shipment to food manufacturers.

Egg white is also air-dried in pan or tunnel systems, in which the egg white is poured in a thin layer on trays. After appropriate treatment of the liquid, as described above, these trays are subjected to temperatures up to 130°F (54°C) until a pseudocrystalline product is obtained. These crystals, or flakes, are used in this form or ground into a fine powder for greater solubility. This type of product is used primarily in the confectionery industry.

Egg yolks. Liquid egg yolk, if not already pasteurized, is held at 140°F (60°C) for 3.5 min. Egg yolk may be spray-dried directly, or it may be subjected to a stabilizing treatment to remove the free glucose, as is also necessary with egg white. The glucose-oxidase enzyme system is used for this process almost exclusively.

Egg yolk must be dried to a final product containing 95% egg yolk solids; many kinds of spray-drying equipment are suitable for this process. Egg yolk is commonly converted to a free-flowing powder by the addition of small quantities of such anticaking agents as sodium silicoaluminate (less than 2%) or silicon dioxide (less than 1%). The ability to flow freely appears to be of particular importance to mayonnaise manufacturers.

Whole egg. Whole-egg solids are manufactured in a similar manner, as are many blends of whole egg, yolk, and added ingredients, to meet specific functional requirements. Common additives are carbohydrate products, which may be derivatives of either sucrose or corn syrup solids. These carbohydrates, when added prior to spray-drying, help retain the original foaming and emulsifying abilities of the liquid egg. Whole egg may be converted to a free-flowing product, as described above for egg yolk.

Frozen products. There is a variety of frozen liquid or solid products that have egg or egg parts as major ingredients.

Liquid. These products include frozen mixes for scrambled eggs, omelets, crepes, and quiches. The scrambled egg mix includes milk products and seasonings and is packaged in waxed paperboard boxes ranging in size from half pint (225 ml) to 2 quarts (1820 ml). A special package primarily for institutional feeding is in a flexible film allowing the eggs to be cooked in the film bag by immersion in hot water at about 194°F (90°C) for times varying with size of the package and temperature at the time of immersion.

Omelet, crepe, and quiche mixes are packed in small, waxed paperboard boxes [half pint (225 ml) to 1 quart (908 ml)]. The package is thawed at 35°F (2°C) and the desired amount poured into an omelet or crepe pan or a quiche shell. After thawing, the mixed product should be kept refrigerated and used within 4 days.

Solid. Solid products include precooked frozen French toast; omelets, either plain or seasoned; crepes, either plain or filled; quiches; souffles; egg patties; fried eggs; diced, hard-cooked eggs; and long eggs.

The French toast is white bread soaked in liquid whole egg and fried to a golden brown color. It is packaged in both retail- and institutional-size packages. Omelets are prepared with whole eggs mixed with about 10% water and seasonings. For variation, other inclusions may be cheese, ham bits, bacon bits, onions, green peppers, or almost any other product desired by the consumer.

Crepes are prepared by using a very thin pancake batter especially rich in eggs. Plain crepes are packaged for retail or institutional use, with the user providing the filler. Crepes are also filled with fruit or meat sauces prior to packaging and freezing. Quiches and souffles are cooked, packed, and frozen for retail distribution or for specialized institutional use such as airline meals.

Other frozen egg products are scrambled egg mix cooked into a patty for sandwich use. Frozen fried eggs are also available for institutional use. The patties and fried eggs are especially suited to microwave reheating at the time of serving.

Hard-cooked diced eggs are usually cryogenically frozen to minimize a toughening of the dices of egg white that can occur when the products are frozen slowly. A unique product is the long egg. This is an extruded product with a center core, about normal egg yolk size in diameter, of cooked egg yolk covered by a uniform layer of hard-cooked egg white. In this product, a small percentage, usually less than 1%, of food gums or starches is added to the albumen to minimize the problem of toughening during freezing.

Refrigerated products. Hard-cooked eggs are prepared by heating shelled eggs to an internal temperature of 175°F (65°C) in hot water. The eggs are then cooled quickly and the shells are removed. The peeled eggs are immersed in 0.5–2% citric acid solution containing about 0.2% of sorbic acid or 0.2% of sodium benzoate as a mold inhibitor. The eggs and solution are stored at 35°F (2°C) and have a microbial shelf life of over 6 weeks.

A modification of the preservative solution including about 5% acetic acid (vinegar) yields pickled eggs. Red beet juice is frequently used with the acetic acid to impart a bright red color to the albumen layer of the hard-cooked egg. Pickled eggs have a refrigerated shelf life of several months.

Another product was introduced in the United States in 1983. These are hard-cooked eggs processed and treated so as to allow assurance of a 3-month usable shelf life. The hard-cooked egg shell is lacquer-coated with a colored solution to seal the egg. Even without refrigeration, these eggs can be kept for 30 days. They are desirable as colored Easter eggs, as ready-to-eat, or for camping trips.

Dry products. The largest use of dried egg products is in bakery and confectionery products. During certain holiday seasons, eggnog mix is popular. This is a mixture of dried milk and dried whole egg with sweeteners and flavoring added. William J. Stadelman

Nonnutritive Sweeteners

Nonnutritive sweeteners are not metabolized to any significant extent and hence do not add caloric value to foods. The U.S. Food and Drug Administration (FDA) defines nonnutritive sweeteners as "substances having less than two per cent of the caloric value of sucrose per equivalent unit of sweetening capacity." Nonnutritive sweeteners include saccharin, aspartame, alitame, cyclamate, acesulfame-K, thaumatin, stevioside, dihydrochalcones, chlorosucroses, L-sugars, and PS-99 and PS-100 (both dipeptide ester compounds). Nonnutritive sweeteners are of particular appeal to those individuals on diets or with diabetes, and to manufacturers interested in producing foods with lower caloric values. Dry foods represent about 30% of the overall market for nonnutritive sweeteners, and liquids account for about 70%. Because of the large economic potential that exists for use of nonnutritive, high-intensity sweeteners in the food and beverage market, a number of materials with such properties are under investigation and development.

Saccharin has been used widely throughout the world for over 80 years. Approximately 60% of the saccharin is consumed in soft drinks, 20% in other beverages and foods, and 20% as a tabletop sweetener. It is about 300 times sweeter than sucrose, is relatively inexpensive, has excellent stability in food systems, and is noncarcinogenic. One disadvantage to saccharin is that it leaves a bitter or metallic aftertaste, particularly when used at relatively high concentrations, which many find objectionable, and which is difficult to mask successfully. In 1977, questions concerning the safety of saccharin, when used in excessively large amounts, were raised by a Canadian animal (rat) study. The FDA decided to ban most of the food uses of saccharin by revoking its approval, except for use by diabetics and the obese, but Congress placed a moratorium on such a ban as a result of public reaction. *See* SACCHARIN.

Aspartame is the methyl ester of a dipeptide derived from aspartic acid and phenylalanine, two common amino acids. It can be produced by chemical, enzymatic, or genetic engineering processes. It is about 180–200 times sweeter than sucrose. Even though aspartame is metabolized, contributing calories, it is considered a nonnutritive sweetener under the FDA definition, since it is more than 50 times sweeter than sucrose. Aspartame has a clean sweet taste with no bitter aftertaste. It is probably the nonnutritive sweetener most closely resembling sucrose in flavor quality and profile. The major problem with aspartame is its relative instability in aqueous solutions, where it hydrolyzes below pH 3 and cyclizes above pH 6. The instability in acid solution delayed its approval for use in soft drinks in the United States. Aspartame was given limited approval for food uses in the United States in August 1981. The approval was for use in dry products such as packets for use as a tabletop sweetener, tablets for hot beverages, instant coffee and tea, cereals, chewing gum, dessert mixes, and powdered drink mixes. In July 1983, approval was extended to carbonated beverages. Aspartame has found widespread use in many of these food products. It is relatively expensive compared to saccharin or sucrose. It has been a controversial sweetener because of questions about its safety, and it must carry a warning label for phenylketonurics. Aspartame has probably been the most studied of all food additives, particularly sweeteners. A patent has been issued for a heat-stabilized composition of aspartame, which could allow its use in baked goods and baking mixes, if approved by the FDA. *See* ASPARTAME.

Acesulfame-K is a crystalline synthetic derivative of acetoacetic acid. It is about 200 times sweeter than sucrose and has excellent stability to heat and low pH. Acesulfame-K was approved in the United States in July 1988 for use in dry goods, such as sugar substitute tablets; chewing gum; dry bases for beverages, instant coffee, and instant tea; dry, free-flowing sugar substitutes in package units; dry bases for gelatins, puddings, and pudding desserts; and dry bases for dairy product analogs. FDA approval will be sought for use in liquids and baked goods. Acesulfame-K is approved for use in 20 countries and is widely used in Europe. It is expensive when compared to saccharin or sucrose.

David R. Lineback

Quality Control

Quality is an important aspect of industrial food production. Quality may be defined as the composite of those characteristics having significance in determining the acceptance of a product unit in relation to the product's design specifications and consumer expectations. Thus, quality control implies a management system for the coordination and integration of all functions of quality such as quality control, quality assurance, and quality reliability necessary to maintain attributes at levels and tolerances acceptable to the consumer while minimizing costs for the producer.

Quality attributes of food products have been divided into three major categories, quantitative, hidden, and sensory. Quantitative attributes of food quality include yield, net weight, drained weight, package integrity, and the like. Hidden attributes of food quality include nutritive value, toxic substances, adulterants, and microbial status. Sensory attributes of food quality encompass those attributes which can be perceived and judged by consumers, and include color, size, shape, consistency, viscosity, texture, mouthfeel, odor, flavor, and defects.

The term food quality control indicates the activities and functions related to controlling the three major categories of food quality during manufacture. In food manufacturing, operating under product quality controls is essential to assure that products meet both company and government (Food and Drug Administration, U.S. Department of Agriculture Food Safety and Inspection Service and the Agricultural Marketing Service, state and local health departments) quality requirements as well as consumer expectations. Controls must be used to keep operational variations in check. This is the function of the quality control department.

A food plant quality-control system should fulfill the following responsibilities:

1. In cooperation with the research and development or procurement departments, or both, written specifications must be prepared for all raw materials, packaging materials, ingredients, supplies, and finished products. Moreover, in cooperation with these departments, written operating procedure specifications must be established for each unit operation of each manufacturing process in the plant.

2. Adherence to sanitation standards and compliance with FDA "Good Manufacturing Practices" must be assured. This includes plant and storage pest control, and plant, storage, and equipment cleanliness including microbiological cleanliness. Employee personal hygiene and work habits should be controlled to assure that foreign material, food safety, and microbiological standards are met.

3. Statistically valid sampling and evaluation procedures must be developed to monitor adherence to ingredient, packaging, raw material, process, unit operation, and finished-product specifications. The results of these evaluations must be recorded on suitable forms and transmitted to management. It is most important that out-of-specification (out-of-control) material be promptly identified and segregated and action be immediately implemented to correct out-of-control conditions, to minimize the production of defective units, and to rapidly resume the production of in-specification product.

Quality-control personnel work in laboratories conducting physical, chemical, and microbiological evaluations of raw materials, ingredients, packaging supplies, work-in-process, and finished product. They also work in the plant to evaluate processes, unit operations, sanitation levels, adherence to "Good Manufacturing Practices," and all other operating specifications. It is in these two functions that considerable reliance is placed upon statistical quality control, whereby statistical techniques are applied to quality evaluations for the purposes of scientific analysis and interpretation of data. Statistical quality-control functions include the selection of appropriate sample size and frequency, the use of attributes and variables sampling techniques and control charts, and the use of correlation, analysis of variance, and evolutionary operations techniques. The methods, procedures, and selection of instruments utilized to measure product quality attributes are all responsibilities of quality control. *See* STATISTICS.

4. In cooperation with the research and development, production, and engineering departments, the quality-control system should evaluate all plant processes for hazard analysis of critical control points (HACCP). The wholesomeness and safety of foods must be assured by identifying and evaluating (using statistical methods) all raw materials and process unit operations in order to preclude the potential for adulteration or microbial contamination which could lead to a consumer health hazard or to inordinate economic loss. It is essential to conduct these evaluations as a safeguard against public health hazards.

5. The quality-control personnel interact cooperatively with regulatory agency personnel to assure that governmental food regulations are understood and followed. Regulations related to processes, "Good Manufacturing Practices," labeling, storage, and shipping of foods must be identified, communicated, and implemented.

6. The system must be able to develop and conduct or assist in an organized program to train supervisors and operators in the quality concept of management. This training has as its ultimate goal the quality consciousness of all plant workers. Additional educational efforts are required for quality-control personnel in order to assure that quality-control workers are versed in basic concepts and perform their duties in a standard, nondeviating manner.

7. Quality control departments should also assist production in improving yields, reducing salvage losses, and improving operating efficiency as well as review and respond to (with production) consumer complaints. Development of a sound recall plan to recover adulterated or defective product from marketing channels is an additional quality-control responsibility. *See* QUALITY CONTROL. Donald V. Schlimme; Barry C. Frey

Food Packaging

Food packaging is a comprehensive system to ensure separation and protection of the contained food product or beverage from external physical, biological, and chemical environmental influences. In synergistically engineering the package material and equipment, consideration must be given to the product contents as well as to the total distribution environment to be experienced during the life cycle of the packaged food product. The objectives of the food packaging system are to provide protection for the products, to effect communications to the consumer, to impart convenience for the consumer, and to create package units containing multiples of individual packages to facilitate distribution through the supply chain.

Regulation. The basic regulations in the United States stem from the 1938 food and drug laws, amended in 1958 to include the Delaney clause on cancer-inducing additives, and since altered to facilitate approvals. The 1964 Truth in Packaging law detailed labeling on the packaging surface. Although much has been written and spoken about packaging laws, these are the two governing laws, with the basic Food and Drug law covering general safety and the Good Manufacturing Regulations covering low-acid foods. Hazard analysis critical control points (HACCP) and nutrition labeling have been added, the former for meats and the latter for all foods.

Labeling. The Truth in Labeling Act stipulates that food packages bear a uniform and accurate representation of the weight, volume, or count of the package contents. The law further provides for accurate word descriptions of the contents, the name and address of the manufacturer or distributor, and the quantity of the contents clearly stated. Not mandated by regulation but by retailers is the requirement that all food packagers contain a universal product code, or UPC.

Indirect additives. The Delaney Amendment has been interpreted to mean that any material that is a potential cancer producer that might migrate from a packaging material in contact with the food in any quantity is regarded as an indirect additive which must be prohibited. The interpretation and implementation of the Delaney Amendment has been manifested in rulings effectively banning poly(vinyl chloride) [PVC] as a liquor packaging material in 1973 and polyacrylonitrile as a carbonated beverage packaging material in 1977. The latter restriction was rescinded in 1984, and the former continues to be under review. The FDA has established limits based on quantities to be ingested plus probabilities of hazard.

Package disposal. The environmental protection movement that began in 1970 has resulted in several states passing legislation that restricts beverage packaging by requiring that packaging for carbonated beverages, beer, and sometimes other beverages bear a deposit. The purpose is to encourage the return of all carbonated beverage and beer containers so that they may be recycled, reused, or visibly removed. A secondary purpose is to encourage the consumer to dispose of the container properly. During the mid-1980s, several states and municipalities enacted laws to encourage recycling of packaging materials, such as funded development programs that forced consumers to separate trash by recyclable and nonrecyclable categories. Some states are attempting to tax packaging. A number of municipalities have passed laws prohibiting certain specific packaging materials, such as expanded (foam) polystyrene. These materials are used more for retail food service than in packaging engineered to protect food from a factory to a store. In addition, "sustainable" packaging—packaging materials derived from renewable resources—is being used. *See* FOAM.

The concern over the potential for upper-atmosphere ozone depletion by fluorocarbons and concurrent global warming has led to legislation prohibiting such materials in packaging such as expanded polystyrene and in aerosol cans. Most manufacturers had already removed the fluorocarbon foaming agents and propellants before the legislation. *See* GLOBAL CLIMATE CHANGE; HALOGENATED HYDROCARBON; OZONE; STRATOSPHERIC OZONE.

Levels of packaging. The levels of packaging usually employed in classification systems are primary, secondary, and tertiary.

Primary packaging is the package in intimate contact with the food product and constitutes the first and usually principal gas–liquid barrier between the food product and its external environment. Primary packaging is generally the metal can, glass jar, plastic jar or bottle, plastic tray or cup, first wrap, inner pouch, and so on into which the food product or beverage is placed prior to being packaged for distribution.

Secondary packaging may consist of a unitizer or container that supplements the function of the primary container. In a unitizer, several primary packages are assembled to permit more efficient distribution. Thus, secondary packaging may include an outer shell paperboard carton, a plastic ring for six cans, plastic shrink film, or multiple packaging for cans or bottles, particularly carbonated beverage or beer cans or bottles, or for soup or seasoning pouches.

Tertiary packaging is the distribution package. Distribution packaging unitizes primary and secondary packaging to facilitate movement through distribution channels such as warehousing and transport. Distribution packaging in the United States is generally a corrugated fiberboard shipping case, kraft paper bundle, plastic shrink film, or some combination. Downstream of tertiary packaging are pallets, pallet stretch film wraps, and so forth.

Materials. A wide range of materials is employed to provide packaging in the United States. The selection of food packaging depends upon such factors as the protection required, economics, and marketing needs. The ability to optimize and combine materials generates multi-ply materials to perform a variety of packaging functions with a single structure.

Metal. Metal cans, usually cylindrical, are rigid; have good vertical compressive strength; resist impact, shock, and vibration; and are able to contain internal pressure or vacuum. To retard the volume decline in metal cans, they are being shaped (such as rectangles, bulbous, undulating, and truncated) to distinguish them on shelves, although shaped cans are more difficult to make and process. Metal cans may be printed directly in the flat before fabrication or in the round, or they may receive an outer paper label to provide the necessary and desired graphics. Metal for food cans has evolved from hot-dipped tin, to electroplated steel, to tin-free steel with chrome–chrome oxide coatings that resist corrosion.

Aluminum is the most widely used material for carbonated beverage and beer cans because of its lightweight and ease of fabrication.

The three-piece metal can (mostly steel) is formed from a flat blank that is wrapped into a cylindrical form. During the 1980s, soldered side seams were replaced by welded side seams to reduce the potential lead contamination hazard. Circular metal disk closures are mechanically clinched to both ends. For consumer convenience, full-panel easy-open can ends increasingly are replacing conventional can ends.

The far more widely used two-piece beverage aluminum can body for carbonated beverages and beer is fabricated from a flat blank or cup. Decoration is usually in the round on the open-ended can. The open-ended can is then transported to the food or beverage packager's line, where it is filled. There the can is closed by mechanical clinching with an aluminum closure. To compete with plastic bottles, narrow-neck aluminum bottles are being offered for both carbonated and noncarbonated beverages. Ends generally may be flat steel for food cans or aluminum with a die-cut to permit easy opening for beverage cans. Diameters of the ends can be reduced by necking-in of can bodies in order to reduce the mass of materials used and to approach bottlelike structures. Finishes may be screw-on/screw-off to more closely match plastic bottles. Although overwhelmingly used for internally pressurized beverages, two-piece aluminum cans are increasingly being used for still beverages such as juices and certain foods by incorporating liquid nitrogen as an internal vacuum compensator to prevent collapse of the light wall aluminum after hot filling and cooling, with subsequent condensation of the internal steam.

Aluminum foil and occasionally steel are also used for nonround containers such as shallow flat trays used for canning. The increase in surface area increases the rate of heat penetration into or out of the package and thus reduces the thermal input required for heat sterilization. Shallow-drawn steel cans with double-seam closures are used for hotel–restaurant–institutional packs. Shallow drawing of both rigid and semirigid aluminum is employed to produce trays in which food is packaged. Aluminum has been generally replaced by plastic, usually crystallized polyester (CPET) and plastic-coated paperboard for frozen and refrigerated foods.

The lightweight and inert properties of aluminum also foster its use as a barrier lamination material in flexible and some semirigid package structures. Pouches and bags employ thin sheets of 0.00028–0.0005 in. (0.0071–0.00127 mm) gage aluminum foil as the principal barrier material in laminations. In lamination form, the aluminum foil is protected against abrasion and pinholing by a coating or laminate, and also has a sealant coating on one face. Because of the relative impermeability of aluminum foil, the metal serves as the principal barrier against the passage of water vapor, oxygen, odors, liquid, other gases, grease, fat, and light.

Vacuum metallization of less expensive substrates, such as plastic film (polyester and oriented polypropylene), often is an alternative. In vacuum metallization, micrometer thicknesses of aluminum metal are deposited on the surfaces of the substrate material to impart significant barrier properties to the substrate. Vacuum metallized surfaces also function as microwave susceptors to convert incident microwave energy into high-temperature heat for surface browning and crisping. Glass (or more accurately, silicon oxide, SiO_x) imparts a gas or water vapor barrier while maintaining both visible-light and microwave transparency. Similar barriers may be affected by aluminum–oxide coatings. *See* PHYSICAL VAPOR DEPOSITION.

Glass. Glass is used for food packaging because of its impermeability to gas, water vapor, and flavor, and its inertness to food products. Glass in bottle or jar form has high vertical compressive strength but relatively low impact strength in the usual configurations and methods of manufacture. Glass is manufactured from very low cost raw materials, but has a high energy input because the raw materials must be melted. Glass bottles may be molded into a wide variety of shapes and sizes to permit the food product to be packaged and distributed in bottles (narrow necks) or jars (wide mouths) that are short or tall.

Reductions in glass weight have been achieved by surface strengthening by methods which include annealing, hot end coatings, and combining with other materials such as roller/dip thin plastic to reduce the potential for shock and vibration shattering and chipping. The latter is used for nonreturnable bottles of carbonated beverages. *See* GLASS.

Closures for glass containers were originally largely metal, either mechanical crown for carbonated beverages or screw-on steel for jars of food products. Major closures for modern glass containers include twist-off crown, roll-on, screw cap, lug cap, tamper-evident/resistant, and snap-ons and press-ons, most of which are available in either metal or plastic with or without internal linings to achieve a barrier with the closure.

Paper. Paper is used for only a small number of packaging applications. Its principal application is as a component of large institutional- and industrial-size multiwall bags for dry granular contents. With high tensile strength and the ability to be coated with barrier plastic materials, unbleached kraft virgin paper is a substrate for manufacture of bags to contain large quantities of sugar, flour, nuts, dry pet foods, and so on. In order to ensure the package integrity, several walls of different materials are employed. These may be combined with other materials such as plastic film or coating to improve the water-vapor barrier properties.

Bleached kraft paper is employed as a laminating material to aluminum foil lamination pouches. The paper acts to protect the aluminum foil from abrasion and puncture, and as a printing surface. Glassine or highly calendered kraft paper is still used as a fat-resistant material, although pearlized plastic, such as oriented polypropylene, is the preferred material. *See* PAPER.

Paperboard. Folding and set-up cartons are primary or secondary packages made from paperboard in a converting process in which the paperboard is printed, scored, and cut. The resulting paperboard blank may then be glued into tubular or tray form to be erected on the packager's line for subsequent filling and closing. Paperboard imparts stiffness, a good printing surface, and some compressive strength. Paperboard made from recycled fibers has been manufactured commercially for more than 100 years and has been used extensively as a secondary packaging material. Recycled paperboard has gradually increased in importance. Usually, a greater mass of recycled material is required as compared to virgin material, because the recycling process reduces fiber length and therefore strength. Recycled paperboard cartons are used on the exterior of bag- or bottle-in-box structures to contain liquid in thin-wall interior plastic film containers, with strength imparted by the exterior paperboard shell. With a water-resistant coating such as polyethylene, virgin solid bleached sulfate paperboard has liquid-resistant characteristics that permit it to be employed as a primary or secondary packaging material for refrigerated beverages or liquid frozen food products. *See* RECYCLING TECHNOLOGY.

Corrugated fiberboard. Usually manufactured from virgin or virgin plus recycled fiberboard, corrugated fiberboard consists of three layers—two outer liners and an inner fluted medium. The fluting or corrugations provide vertical stacking strength in the direction of the flutes and cushioning against vibration and impact damage. Corrugated fiberboard sheets are printed, cut, and formed into cases; or in the reverse process, are printed after fabrication. An advance in corrugated fiberboard case manufacture has been the employment of rubber- or plastic-plate flexographic printing processes. Flexographic printing permits corrugated fiberboard to be used not only for distribution packaging but also for display and consumer packaging. Increasing quantities of corrugated fiberboard case liners are being laminated to offset printed paper to provide superior graphics and the basis for consumer-oriented cartons. Increasing numbers of corrugated fiberboard cases are being made with white outer linerboard which provides a printing surface superior to traditional brown outer liner. Often such outer linerboard is preprinted by highly effective processes to deliver excellent graphics. Improved compression strength in final cases is achieved by advanced papermaking such as cross dispersion of fibers on the Foudrinier wire, by multiple laydown of fibers, or by off-line water-resistant coating techniques.

Coated and laminated paperboard. Paperboard that is extrusion-coated with polyethylene on two sides or laminated with polyethylene to aluminum foil and further extrusion-coated with polyethylene can be fabricated into liquid-tight containers. Laminated materials are being used for aseptically packaged and hot-filled/refrigerated–distribution beverages. Polypropylene as an inner coating acts as a flavor barrier. Ethylene–vinyl alcohol copolymer coating is an oxygen barrier and a paperboard flavor barrier. A wide variety of polyethylene-coated paperboard liquid containers are employed for milk, dairy drinks, and fruit beverages. Lightweight liquid-tight containers with water-repellent and water-resistant coatings are employed for packaging dairy and delicatessen items in retail operations. These are frequently fabricated into cups, tubs, or trays from virgin bleached sulfate or sulfite paperboard, coated with polyethylene.

Composite paperboard cans. Kraft paper and recycled paper are laminated to other materials which are then fabricated by spiral winding into multi-ply paperboard-body composite cans with metal, foil lamination, or paper ends. Spiral-wound composite cans have the same shape as metal cans but have lower barrier characteristics. Spiral-wound composite cans contain frozen concentrated orange juice, refrigerated doughs, snacks, and granular dry beverage products. Spiral-wound composite paperboard cans are closed by heat-sealing aluminum foil laminations to the flange rather than by traditional mechanical seaming with rigid metal. To some degree, convolute-wound single-wall cans may be used for some applications.

Plastics. Plastic materials are made from organic resins or polymers. Nearly all plastics used for food packaging are thermoplastic and thus are sealable by heat and pressure. A small quantity of thermoset plastics is used for bottle closures. Thermoplastics are favored for packaging because of their light weight.

Cellulosic films. The first of the so-called plastic materials were derived from regenerated or otherwise chemically altered cellulose materials. Commercially, cellophane was the most important.

Cellophane is a transparent film composed of a plasticized base sheet of regenerated cellulose. Essentially greaseproof and coated for moisture barrier, cellophane was employed in applications such as primary packaging for snacks, overwraps on cartons, bags, pouches, windows on cartons, and as a component of lamination. Because of the availability of more functional and less expensive plastic materials, the use of cellophane as a packaging material has declined dramatically. *See* CELLOPHANE.

Polyethylene. The major plastic employed for food packaging is polyethylene. Low-density polyethylene with a density range between 0.915 and 0.925 g/cm^3 is largely used for making film, with the range of film gages 0.00075 in. to about 0.005 in. (0.02 to 0.127 mm). Very low density polyethylene resins have desirable properties, such as toughness and heat seal. Low-density polyethylene is a very tough, water- and water-vapor–resistant film. Gas barrier characteristics of all of the polyolefin films are not good.

Because of polyethylene's effectiveness as a water vapor barrier, it is frequently used as a bagging material to contain water vapor–sensitive products. Low-density polyethylene film is a component of a multiwall bag system with paper as another major component. It is also used as a sole component of large bags. Low-density polyethylene is used in soft baked goods bags in which short-term water vapor resistance is required.

Low-density polyethylene resin also is used extensively as an extrusion coating resin on paperboard and aluminum foil to impart both water vapor resistance and heat sealability. Polyethylene resin imparts liquid tightness to milk and juice cartons.

Linear low-density polyethylene, produced by lower-pressure processes, has been replacing low-density polyethylene. Linear low-density polyethylene permits broader heat-seal ranges and greater toughness, so that thinner films may be blown or more durable heat seals on pouches may be made.

High-density polyethylene film. High-density polyethylene is a plastic resin with density above 0.93 g/cm^3. The resulting material in film or other form has an excellent water vapor resistance but, like all polyolefins, relatively poor gas transmission. High-density polyethylene film is stiffer than low-density polyethylene and has a relatively high operating temperature. In film form, this material is used to provide high-temperature properties to paper, paperboard, or aluminum foil. It is also used for injection and blow molding to produce semirigid containers such as milk and juice bottles, jars, trays, and tubs. Many food products are packaged in high-density polyethylene bottles, provided there are no high-temperature or special oxygen barrier requirements. A tougher version, high-molecular-weight high-density polyethylene, has permitted significant down-gaging of film in bag and can or drum liner applications. With the majority of semirigid bottles and jars being manufactured from high-density polyethylene, and with increasing recovery of so-called used-bottle infrastructure, this material has become a dominant plastic in food packaging.

Polypropylene. Polypropylene film is formed by casting through a slot extruder or, less frequently, by blowing through a circular die. Cast polypropylene, however, is not readily heat-sealable and is low-temperature-sensitive. If the film is biaxially stretched and heat-set, it acquires a much improved low-temperature impact resistance and strength. With appropriate coatings such as poly(vinylidene chloride) or acrylics from coextrusion or off-line coating, oriented polypropylene (OPP) film is employed for bag and pouch making, and overwrapping. Oriented polypropylene may be printed by the methods employed for polyethylene, and so its properties as a film in common gages around 0.001 in. (0.025 mm) are satisfactory for packaging of snacks, candy, bread, and baked goods.

Polypropylene has a relatively high operating temperature, 255°F (124°C), and can be employed as a sealing medium on laminations in which high-temperature resistance is required, such as retort pouches. Polypropylene is also employed as a semirigid bottle- or tray-molding material structural component in selected applications such as retort trays or hot-filled catsup. *See* POLYOLEFIN RESINS.

Poly(vinyl chloride). Although it has been under adverse pressure by environmentalists because of its chlorine content, poly(vinyl chloride) [PVC] is used in highly plasticized form for film. The film application is to permit oxygen entry to retain the red color or bloom on the surface of fresh meat packaged in retail store backrooms.

Poly(vinylidene chloride). Poly(vinylidene chloride) [PVDC], or saran, is an outstanding oxygen, water vapor, grease, and liquid barrier. Its price, however, is high, limiting its use to very thin films or coatings to provide a water vapor and gas barrier on the surfaces of other plastic or paperboard substrates. PVDC may be coextruded with other materials to provide a high gas barrier. PVCD is copolymerized with other monomers to enhance barrier properties. *See* POLYVINYL RESINS.

Ethylene–vinyl alcohol. Ethylene–vinyl alcohol (EVOH) resins are high-gas-barrier materials that are used in coextrusion and coating for both semirigid and film applications. Generally, these resins have moisture sensitivity that increases with decreasing ethylene content or higher gas-barrier properties.

Polyester. Poly(ethylene terephthalate), also known as PET or polyester, is a high-strength family of plastic materials with outstanding dimensional stability, good transparency, good heat resistance, and modest gas and water vapor barrier properties. Injection-stretch blow-molded polyester carbonated beverage bottles are used commercially despite having only intermediate carbonation retention and oxygen hold-out characteristics. The bottle is the most rapidly accepted new package in United States packaging history. Polyester is used for packaging a variety of other bottled food products with intermediate oxygen-barrier requirements, including peanut butter and salad dressing, as well as water.

Polyester is the basis for dual ovenable (microwave or convection oven) packaging in both thermoformed forms and as extrusion coatings on paperboard. *See* POLYESTER RESINS.

Nylon (polyamide). Nylon materials are good gas and aroma barriers which are highly thermoformable. Although somewhat moisture-sensitive, they are widely used in processed meat and cheese packaging as the forming web. Amorphous nylon and other variations were proposed as a blow-molding material in high-oxygen-barrier jars and bottles but were not widely adopted. *See* POLYAMIDE RESINS.

Laminations. Although single-layer materials have been employed for packaging for many years, laminations represent a major step forward in packaging. Lamination is defined as producing film materials independently and then assembling or bonding two or more films or paper or aluminum foil together. In this manner, various functional materials can be combined to produce a single flexible web with the desired functional properties and usually at a lower cost than a single material alone.

Two basic methods of lamination are extrusion and adhesion lamination. In the former, thermoplastic material is applied to a substrate and employed as an adhesive that holds together two materials such as paper and aluminum foil. The alternate method is to apply a thin film of adhesive and then bond the adhesive layer to the second layer.

Coextruded films. Coextrusion unifies thermoplastic melts in a single extension die to intimately bond optimum quantities of plastic materials and form either sheets or film. Coextrusion often requires intermediate extrudable adhesive tie layers.

Coextruded films are used extensively for snack-food pouch packaging. Polypropylene/polyethylene coextrusions are employed for candy bar wrapping and snack food pouches.

Coextruded sheet is employed for multicolor and surface-gloss tub packaging of dairy products. Coextruded sheet containing oxygen barrier materials, such as PVDC or ethylene–vinyl alcohol copolymer, are used for packaging highly oxygen-sensitive products such as processed meats and cheeses. Plastic beer bottles are being fabricated in multiple layers by coinjection of polyester and a physical (other polymers) oxygen barrier or a chemically active layer (polymer containing oxygen scavengers).

Other comaterial plastics. Other methods of producing layers of materials include extrusion followed by blow or injection molding, extrusion followed by thermoforming and blowing, or variations thereof. The resulting products can be bottles or jars containing several layers of materials, one of which is a structural material, a second of which might be a water vapor barrier, and the third perhaps an oxygen barrier. These high-barrier bottles have found application for packaging oxygen-sensitive products such as ketchup. *See* PLASTICS PROCESSING.

Active packaging. Active packaging senses changes in the internal or external package environment and adjusts the packaging function to the change. Among the more prominent active packaging systems are moisture control, oxygen scavenging, carbon dioxide generation, gas permeability control, and antimicrobial inhibition. Intelligent packaging senses change signals such as maximum temperature, time–temperature integration, location (radio frequency identification, or RFID), or spoilage signalers. Aaron L. Brody

Package selection. The selection of packages appropriate for a food product involves decisions that satisfy a matrix of requirements. The matrix is the combination of the functions of the package and the environments in which it must perform.

Functions. The functions of a package are containment, protection, convenience, and communication. Containment simply means that the package allows the product to be handled and transported. Packages protect food products from spoilage agents, which include vermin, microorganisms, dust, mechanical forces, odors, moisture (addition or removal), and oxygen. Thus, packaging prolongs the shelf life of the product, that is, the length of time during which the product retains an acceptable quality level. Convenience, a function that frequently drives the development of new types of packages, has two aspects: apportionment, and handling and storage. Apportionment involves the selection of package capacities convenient for the consumer or ultimate user. Cheese packaged in 1-lb-capacity (0.5-kg) pouches is much more convenient for consumer usage than cheese available in 50-lb (23-kg) wheels.

The communication function of a package involves the transfer of information and the selling of the product in the package. The label on a food product provides quantifiable information concerning grade, quantity, ingredients, and nutrition. This is a transfer of information. The label also may have a picture of the product or some other image meant to convince the consumer to purchase.

Performance environments. The environments in which food packages must function are the production (food processing) environment, the distribution environment, and the consumer or other ultimate user environment, including the disposal of the package. The properties of a food package that allow its selection for use with a particular product begin with the process requirements of the food product. If the food product is to be retorted or pasteurized in the package, the package must remain inert through this high-temperature processing. The package must also be suited to the equipment available for filling, closing, labeling, and other operations at the food processor at an acceptable cost. The distribution environment (including transport, warehousing, and often display) introduces requirements to protect against physical damage and ambient (temperature, humidity, gas permeation) spoilage. Physical damage to a package (for example, a dented can) or to a product (such as a crushed snack food) is due to impact, compression, and vibration. Impact damage occurs when a package is dropped or, in a railcar, when cars are coupled together. Compression damage generally occurs in warehouse storage or in transport, when heavy loads are placed upon a package, often by stacking pallets. Vibration damage occurs during transport because of the vibration of a truck, train, ship, or airplane. In vibration damage, packages or products are rubbed together and damaged by abrasion. Ambient conditions, such as high moisture or low oxygen, similarly cause food spoilage, either independently or in combination with physical abuse. High moisture levels in a warehouse, for example, diminish the strength of corrugated boxes, allowing the bottom boxes on a pallet to be crushed by those above. Oxygen in air causes rancidity spoilage of fats and oils. These foods need to be protected from the oxygen in the ambient environment. Other foods, such as fresh vegetables, require available oxygen for respiration. Holes are cut in their packages to allow oxygen availability.

Typical examples. The selection of a package for a particular food product begins with knowledge of the properties of the food relevant to maintenance of its quality. The pH of the food, its susceptibility to microbial spoilage, its interactions with oxygen and moisture, and the chemical reactions diminishing quality need to be considered. Thereby, some packaging materials might be ruled out. Paper and other oxygen-permeable materials would be ruled out for oxygen-sensitive foods. Food processing steps after packaging (pasteurizing, retorting) will rule out some other materials.

The selection of the appropriate polymer for flexible packaging of a particular food generally requires thorough analysis. The primary differences among resins concern cost, strength, barrier, haze, and heat sealability. The simple commodity resins (polyethylene, polypropylene) will have the lowest cost per pound, the complex barrier resins (EVOH and PVDC) the highest. Strength values between the weakest and strongest resins cover an order of magnitude; oxygen barrier values differ by 5 orders of magnitude. A slight difference in haze (1% versus 3%) is noticeable to the consumer. For each type of resin, suppliers offer several grades, each with its own combination of properties. Often, a particular grade will perform best on a particular packaging line.

Peter J. Vergano

Nutrition Labeling

The nutrition label provides consumers with standardized information on the primary nutritional characteristics of food products. The nutrition label can serve as a guide in the planning of a healthy diet. In the United States, the Nutritional Labeling and Education Act of 1990 resulted in the formulation of new nutritional labeling regulations by the Food and Drug Administration and the Department of Agriculture. Nutrition labels are required on most packaged foods. For some bulk foods at retail, a nutrition label must be displayed at the point of purchase. The use of nutrition labels is voluntary for some foods, including raw meats, poultry, fish, fruits, and vegetables.

The nutrition label in the United States is headed by the words "Nutrition Facts." The serving size, stated in both the metric system and the U.S. Customary System, and the number of servings per container are provided. The serving sizes for many foods were established in the federal regulations and must be used for the calculation of the number of servings per container. Calories per serving and the number of calories from fat per serving are provided next. Dietary guidelines suggest that consumers obtain no more than 30% of calories from fat, and the percentage can easily be calculated from these two pieces of information for each food.

A list of the nutrients and other dietary components most important to health follows. The list includes total fat, saturated fat, cholesterol, sodium, total carbohydrate, dietary fiber, sugars, and protein. The amount of each per serving (in metric units) and the percent daily value provided by a single serving for each are given on the label, with the exception of sugars and protein where only the amount per serving is provided. Certain other components of the food can be listed on a voluntary basis, including polyunsaturated fat, monounsaturated fat, potassium, soluble and insoluble fiber, sugar, alcohols, and other carbohydrates.

The percent daily value, as derived from the daily reference value (DRV), is particularly helpful in devising healthful eating strategies. On the lower part of the nutrition label, the DRVs for fat, saturated fat, cholesterol, sodium, total carbohydrate, and fiber are provided in metric weight units for two typical human daily diets, a 2000-calorie diet and a 2500-calorie diet. These amounts are established by regulation and are consistently provided on all packaged foods bearing these nutrition labels. For example, the DRVs for a 2000-calorie diet are 65 g for total fat, 20 g for saturated fat, 300 mg for cholesterol, 2400 mg for sodium, 300 g for total carbohydrate, and 25 g for dietary fiber. For total fat, saturated fat, cholesterol, and sodium, these DRVs would represent the maximum amounts that are desirable in a 2000-calorie diet. The percent daily value that appears on the label adjacent to each nutrient or other component is calculated by dividing the amount of that component by the DRV of that component for a 2000-calorie diet. For example, a food product with 10 g of saturated fat per serving would contain 50% of the DRV for saturated fat on the basis of a 2000-calorie diet. To calculate appropriate healthful eating strategies, individual consumers should adjust the DRVs to fit their own ideal caloric intakes and calculate the percent daily value on that adjusted basis. Calculation of the percent daily value as adjusted for the consumer's ideal caloric intake will show how each individual food fits into the overall daily diet. Food choices can also be easily compared for their amounts and contributions to the DRV on a per-serving basis. Such comparisons allow consumers to make intelligent choices about the composition of their diets based on their own individual needs.

Following the list of major nutrients and other dietary components is a listing of several of the most important micronutrients in the diet—vitamin A, vitamin C, calcium, and iron. Their levels are declared as a percent daily value based upon the Reference Daily Intake (RDI), which is identical to the U.S. Recommended Dietary Allowances (RDA) as established for adults and children over the age of 4. The RDA values were established by the Food and Nutrition Board of the National Academy of Sciences, and they represent an amount of the micronutrient that will be sufficient to prevent any signs of deficiency diseases. Other vitamins and minerals can be listed voluntarily. However, their listing is required in the case of intentionally fortified food products, where a percent daily value must be given also.

The nutrition label provides, at the bottom, the comparative number of calories per gram provided by fat (say, 9 calories per gram), carbohydrates (4 calories per gram), and protein (4 calories per gram). This information explains how calories per serving can be calculated by knowing the amounts of fat, carbohydrates, and protein per serving.

Most packaged foods display nutrition labels containing all of the information described above. In a few cases, simplified labels are allowed. Examples include foods in small packages and foods with insignificant levels of some components, for example, soft drinks.

Steven L. Taylor

Bibliography. E. D. Aberle et al., *Principles of Meat Science*, 4th ed., 2001; M. Ash and I. Ash (eds.), *Handbook of Food Additives*, 1995; A. Brody, *Active Packaging Beyond Barriers*, 2002; A. Brody and J. Lord, *Developing New Food Products for a Changing Marketplace*, 2000; A. Brody and K. Marsh, *The Wiley Encyclopedia of Packaging Technology*, 2d ed., 1997; R. P. Carpenter, D. H. Lyon, and T. A. Hasdell, *Guidelines for Sensory Analysis in Food Product Development and Quality Control*, 2d ed., 2000; H. Charley, *Foods: A Scientific Approach*, 3d ed., 1997; F. Clydesdale and F. J. Francis, *Food, Nutrition, and Health*, 1985; J. J. Connell (ed.), *Control of Fish Quality*, 4th ed., 1995; P. M. Davidson et al. (eds.), *Antimicrobials in Foods*, 3d ed., 2005; R. B. Fast and E. F. Caldwell (eds.), *Breakfast Cereals and How They Are Made*, 1990; Food and Drug Administration (U.S.), *Food Labeling Guide*, 1999; W. A. Gould, *Tomato Production, Processing and Technology*, 1992; W. A. Gould and R. W. Gould, *Total Quality Assurance for the Food Industries*, 3d ed., 2001; R. C. Griffin, S. Sacharow, and A. L. Brody, *Principles of Package Development*, 2d ed., 1993; B. Holland (ed.), *Fish and Fish Products*, 1994; J. M. Jay, M. J. Loessner, and D. A. Golden, *Modern Food Microbiology*, 7th ed., 2005; W. A. Jenkins and J. P. Harrington, *Packaging Foods with Plastics*, 1991; J. M. Jones, *Food Safety*, 1992; M. J. Kirwan (ed.), *Paper and Paperboard Packaging Technology*, 2005; A. Lopez, *A Complete Course in Canning*, Books 1 and 2: *Basic Information on Canning*, 12th rev. ed., 3 vols., 1987; M. McWilliams, *Food Fundamentals*, 8th ed., 2006; S. A. Matz, *Chemistry and Technology of Cereals as Food and Feed*, 2d ed., 1991; S. A. Matz, *Snack Food Technology*, 1992; M. Mathlouthi, *Food Packaging and Preservation*, 1994; N. H. Mermelstein, A new era in food labeling, *Food Technol.*, 47(2):81–86, 94, 96., 1993; N. N. Potter and J. H. Hotchkiss, *Food Science*, 5th ed., 1995; H. Reuter, *Aseptic Packaging of Food*, 1989; G. L. Robertson, *Food Packaging Principles and Practice*, 2d ed., 2005; D. B. Sauer (ed.), *Storage of Cereal Grains and Their Products*, 4th ed., 1992; S. Selke, *Understanding Plastics Packaging*, 1997; R. P. Singh and D. R. Heldman, *Introduction to Food Engineering*, 3d ed., 2001; J. Smith (ed.), *The Technology of Reduced-Additive Foods*, 2d ed., 2004; W. J. Sultan, *Practical Baking*, 5th ed., 1990; R. T. Toledo, *Fundamentals of Food Process Engineering*, 3d ed., 2006; D. H. Watson (ed.), *Food Chemical Safety: Contaminants*, 2001; D. H. Watson (ed.), *Food Chemical Safety*, vol. II: *Additives*, 2002; R. B. H. Wills, *Postharvest: An Introduction to the Physiology and Handling of Fruit and Vegetables*, 4th ed., 1998.

Food microbiology

A subdiscipline in the field of microbiology concerned with the study of bacteria, fungi, and viruses that grow in or are transmitted by foods. While bacteria are frequently associated with food spoilage and food poisoning, some species preserve foods through fermentation or produce food ingredients. Food microbiology is a broad field that can include not only microbiology but also sanitation, epidemiology, biochemistry, engineering, statistics, and mathematical modeling.

Pathogens and spoilage organisms. Some people dismiss food poisoning as a minor annoyance. In reality, the suffering and economic losses stemming from food-borne pathogens are substantial, but they are often hidden. Annual economic losses from food-borne pathogens are extremely high. *Salmonella*, which cause an average of 40,000 cases yearly and 2000–3000 deaths in the United States, are responsible for about a third of these losses. Individual outbreaks of food-borne diseases can affect thousands of people. Many outbreaks are predictable and preventable through good sanitation, preservatives, thermal processing, and refrigeration. More than half, however, are of unknown etiology are poorly understood, and may be caused by so-called new pathogens. *See* FOOD POISONING.

Historical pathogens. In the 1960s, most food-related illnesses were attributed to one of five major groups of pathogenic bacteria. These were associated with particular foods, commodities, or processes and were classified as infectious or toxin-producing. These five groups, described below, remain major causes of food-borne illness. *See* BACTERIA.

Salmonella and Shigella. The primary infectious bacterium associated with foods is *Salmonella*. These organisms cause gastroenteritis with symptoms of fever, diarrhea, and vomiting 12–36 h after ingestion. Salmonellosis is usually self-limiting, but it can be fatal in the old, young, or medically compromised individuals. *Salmonella* are commonly found on meats, especially poultry and eggs. *Salmonella* are easily killed by cooking. However, items contacted by the contaminated raw meat can transfer the *Salmonella* to food that is ready to eat (cross contamination) and cause illness. The seasonal increase in *Salmonella* isolations illustrates how food-borne illness increases in warm summer months.

Shigella are related organisms which produce a similar infectious syndrome. They are usually transmitted by a fecal-oral route or through feces-contaminated water rather than through foods.

Clostridium botulinum. The most dreaded toxin-producing organism is *Clostridium botulinum*. It excretes a potent neurotoxin that causes weakness, double vision, slurred speech, paralysis, and often death if ingested. The vegetative reproductive form of *C. botulinum* is heat-sensitive, lives only in the absence of air, does not compete well with other bacteria, and is rarely a problem in fresh foods. *Clostridium botulinum* spores are killed only through severe heating, such as in canning.

Historically, botulism has been associated with foods canned at home. If canned foods receive inadequate heat processing, competing bacteria are killed, air is expelled, and the botulinal spores germinate. Fortunately, botulinal toxin is often destroyed by heat when the food is cooked before serving; hence the standard advice is to boil home-canned foods before eating.

Modern commercial canning is designed to destroy *C. botulinum* spores. Reported outbreaks of botulism caused by pot pies, potatoes, and fried onions have been caused by temperature abuse, that is, the holding of foods at warm temperatures that promote bacterial growth. *Clostridium botulinum* can also be a problem in processed meats, such as hams and sausages. In this case, its growth is controlled through the use of nitrite, salt, and refrigeration. One type of *C. botulinum* is associated with fish. *See* BOTULISM; TOXIN.

Clostridium perfringens and Bacillus cereus. These are spore-forming, toxin-producing bacteria that cause illness when foods are heated enough to kill competing bacteria but not enough to kill the spores. When large volumes of foods are prepared, cooked, and then kept warm until they are served, spores can germinate. In the case of *C. perfringens*, which is associated with meats, the ingested cells release toxin in the digestive tract, resulting in cramps and diarrhea. *Bacillus cereus*, found in meats, dried foods, and rice, produces two different types of toxins: the diarrheal toxin, which has an etiology similar to *C. perfringens,* and the emetic (vomiting) toxin, which causes symptoms similar to those produced by staphylococcal toxins.

Staphylococcus aureus. This bacterium produces toxins that are very resistant to heat. *Staphylococcus aureus* is found in the nose and throat of many healthy people and is transferred to food by inadequate hygiene. When foods are temperature-abused, the bacteria grow and produce toxin. Subsequent heating of the food kills the bacteria but does not inactivate the toxin. The toxin causes severe vomiting and diarrhea from $^1/_2$ to 4 h after ingestion. The microorganism grows well at salt and sugar concentrations that inhibit many competing bacteria. Foods high in protein, such as cured meats, custards, and cream-filled bakery goods, pose special hazards for staphylococcal food poisoning. *See* STAPHYLOCOCCUS.

Microbial ecology of foods. Modern food microbiology views foods as habitats where different organisms compete for survival. The fact that there are 250 genera of bacteria and that only 25 of these (8 pathogenic) are found in foods suggests that foods provide unique ecological niches. Viruses do not reproduce in foods and are not competitors in this sense (the food acts only as a carrier). Yeasts and molds usually grow more slowly than bacteria and are rarely a problem in foods that support bacterial growth. *See* FUNGI; YEAST.

Bacteria reproduce by binary fission; it takes only 20 doublings for one cell to yield more than 1 million cells. In environments where the doubling time is short, this occurs quite rapidly. Many preservation methods alter foods' environmental conditions in order to slow microbial growth. *See* BACTERIAL GROWTH.

Temperature. The most important environmental condition is temperature. Most food-borne pathogens are mesophiles; that is, body temperature is optimal for growth. With a doubling time of 20 min at 98.6°F (37°C), one bacterium generates 1 million progeny in less than 7 h; at 32°F (0°C) the doubling time increases to 1200 min and the 1 million cell count is not reached for 16 days. Keeping hot foods hot (>145°F or 63°C) and cold foods cold (<45°F or 7°C), combined with rapid heating and cooling to get rapidly beyond the growth-promoting temperature range (45–145°F or 7–63°C), prevents most foodborne illnesses. Psychrophylic (cold-loving) bacteria such as *Listeria monocytogenes* are exceptions.

Acidity. A food's acidity, quantified as pH, is another major environmental factor. The pH range for bacterial growth is 4–9, with fastest growth at neutrality (pH 7). Changing a food's acidity can change the rate of bacterial growth. Meats, fish, poultry, and most dairy products are near pH 7, which is ideal for bacterial growth; fermented foods and fruits have pH less than 4. Many yeasts and molds grow in acidic environments and spoil acidic foods. The pH value of 4.6 has special significance because *C. botulinum* can grow and produce toxin above this value. Canned foods with pH above 4.6 are legally classified as low-acid and must be processed in retorts under steam at 240–280°F (116–138°C) to kill *C. botulinum* spores. Foods with pH below 4.6 are legally high-acid and are processed in open pans of boiling water. In this case, *C. botulinum* need not be killed because it cannot grow at low pH. *See* PH.

Water activity. The amount of water available for microbial growth, that is, water activity (a_w), is the third major factor influencing microbial competition. Water activity is the equilibrium relative humidity generated by a food in a closed chamber divided by 100 to give a 0 to 1.00 scale. Salad dressings and honey, which both contain 50% water, are microbiologically quite different. The dressing separates into a 100% free-water phase ($a_w = 1$) and supports bacterial growth, while the sugar in honey binds water so tightly that it is unavailable for microbial growth. Most bacteria grow only at $a_w = 0.90$–1.00. Fresh meats, vegetables, fruits, and perishable foods have water activity in this range. Most yeasts can grow at slightly lower values. *Staphylococcus aureus* is the pathogen most insensitive to water; it grows at $a_w = 0.86$. Since no pathogenic bacteria grow below $a_w = 0.85$, this value has special significance in the regulations defining low-acid foods. Foods having an a_w value below 0.85 are legally considered high-acid, regardless of their pH. Most molds grow at a_w values as low as 0.8 and compete well in foods such as flour, cakes, beans, rice, and cereals. Some xerophilic molds and yeasts grow at a_w values as low as 0.6. Dehydrated foods, with even less available water, are completely recalcitrant to microbial spoilage.

Oxygen. Oxygen can be favorable, neutral, or inhibitory to bacterial growth. In one process, foods are first vacuum-packed to inhibit aerobic spoilage organisms and are then partially cooked. This environment is perfect for anaerobic spore-forming microorganisms. However, the Food and Drug Administration prohibits the use of this process because of the potential botulinal hazard.

Preservatives. Chemical preservatives also render food environments unsuitable for microbial growth. The oldest preservative is common table salt; at very high levels, it produces water activities that are inhibitory to microbial growth, although many organisms are inhibited by as little as 3% salt. Nitrites are used in cured meat as anticlostridial agents. Acetic, lactic, citric, benzoic, and propionic acids and sodium diacetate can also be added to foods as microbial inhibitors. Considering the low levels used, the long history of safe use, and the consequences of microbial growth, the risk:benefit ratio associated with chemical preservatives is very low.

Multiple barriers. Consumer preferences for fresh and natural foods make it difficult to alter any one environmental factor enough to inhibit microbial growth, and so the trend is to use multiple barriers, or hurdles. This approach employs several inhibitors at suboptimal levels. For example, clostridia may be inhibited by 7% salt at pH 7.0 or 0% salt at pH 4.6, but a meat treated this way would be unacceptable, tasting either salty or acidic. However, 3% salt at pH 6 in the presence of nitrite at a concentration of 125 parts per million provides multiple barriers sufficient to inhibit the bacteria and not impair flavor.

Emerging pathogens. The demand for longer shelf life in refrigerated foods combined with their increased popularity has caused renewed concern about psychrophilic pathogens, such as *Yersinia enterocolitica*, and enterotoxigenic organisms, such as *Escherichia coli* and *Listeria monocytogens*. These bacteria grow most rapidly at 59–86°F (15–30°C) and, at refrigerated temperatures, can succesfully compete with the normal mesophilic bacteria, thus limiting the shelf life of refrigerated foods. *Listeria* cause special concern because they infect women and their unborn children preferentially. *See* ESCHERICHIA; LISTERIOSIS; YERSINIA.

Campylobacter jejuni is a pathogen that is responsible for more illnesses than *Salmonella* and *Shigella* combined. Ingestion of relatively small numbers can cause diarrhea, cramps, and nausea. This organism is microaerophilic (requires 5–10% oxygen), is relatively fragile, and shows a seasonal pattern of outbreaks similar to *Salmonella*. It is associated with raw meats and unpasteurized milk, and can be controlled by pasteurization, heating, and good sanitation.

Analytic approaches. Microbial analysis of foods frequently requires "zero defects" in the absence of 100% testing. Legally, ready-to-eat foods must be free of *Salmonella*. This demands that the food microbiologist be able to detect one *Salmonella* among millions of innocuous bacteria in a pound of food. Moreover, all of the food cannot be tested because microbial analysis is destructive. Therefore, statistical sampling plans determine how many samples must be tested to have confidence that the whole lot is free of *Salmonella*.

In the classical methods for counting microorganisms, a food or its hemogenate is highly diluted so that only 30–300 cells are transferred to growth media. After 2–10 days, each cell grows into a colony, and these are counted and multiplied by the dilution factor to estimate the number of cells in the food. Automated methods have been developed that measure growth products, bacterial deoxyribonucleic acid (DNA), or specific toxins; these methods dramatically reduce the analysis time and are rapidly replacing the petri-dish method.

A procedure known as hazard analysis critical control points (HACCP) can replace much postproduction testing. This technique examines a food, its ingredients, and its processing to identify points critical to safety. These points are then heavily monitored during production; if they are maintained, a safe product results.

Beneficial food-borne organism. When certain bacteria grow in foods, they produce desirable flavors and textures, and may also inhibit pathogenic organisms. Most of these bacteria belong to the genera *Streptococcus*, *Lactobacillus*, *Leuconostoc*, *Pediococcus*, or *Micrococcus*. They are used to make fermented dairy products, meats, and vegetables, and to preserve food by converting the sugars needed by competing microbes to lactic acid, which inhibits their growth. These lactic acid bacteria are unusually tolerant of acidic environments. *Acetobacter* and *Gluconobacter* are used in the production of vinegar. Yeasts, usually *Saccharomyces*, which produce ethanol and carbon dioxide, are used in the processes of brewing and baking.

Lactid acid coagulates casein, the major protein in milk, and this process is used to manufacture cheese. During the aging of cheese, bacterial enzymes generate characteristic flavors, allowing a wide variety of products to be made by using many different bacteria. The bacteria used can be indigenous to the milk, derived from a previous fermentation, or added as pure cultures. *See* CHEESE; FERMENTATION.

Until the late 1960s, staphylococcal food poisoning was a major problem in certain meat products, such as bologna, pepperoni, and salami. Since acid production by indigenous bacteria is often unpredictable, it is now recommended that defined starter cultures be used to ensure that sufficient acid is produced early enough to prevent staphylococcal growth.

A novel use for starter cultures of lactic acid bacteria is to prevent botulinal growth in bacon. A small amount of culture and sugar are added to the cured meat; if the bacon is temperature-abused, the lactic acid bacteria grow and produce acid to inhibit botulinal growth.

Certain vegetables are preserved by fermentation. Pickles are made by fermenting cucumbers; olives and many oriental foods are also fermented. During sauerkraut fermentation, the addition of 2.5% salt (by weight) to shredded cabbage selects for the growth of *Leuconostoc mesenteroides*, which stop growing when acid levels reach about 0.067%. This environment favors *Lactobacillus plantarum*, which produces acid to levels of 1.25%, which is tolerated by *L. brevis*, and this bacterium brings the product to a final acidity of 1.7%.

Biotechnology. Advances in molecular biology have generated interest in applications to food processing.

The most important contribution of biotechnology to food microbiology is the production of probes that detect pathogenic organisms much faster than conventional methods. For example, conventional methods require 5 days to confirm the presence of *Salmonella* in foods; probes that detect *Salmonella*-specific DNA or antigens can give similar results in 2 days.

The dairy industry has benefitted from advances in biotechnology by acquiring the ability to determine the genetic basis for the bacterial metabolism of lactose in milk and to stabilize it. In addition, enzymes that accelerate the aging of cheese have become commercially available, making it possible to produce a cheese with the taste of 9-month-old cheddar in just 3 months. *See* BIOTECHNOLOGY; ENZYME; FOOD ENGINEERING; FOOD MANUFACTURING; FOOD PRESERVATION; GENETIC ENGINEERING; VIRUS.

Thomas J. Montville

Bibliography. W. C. Frazier and D. C. Westhoff, *Food Microbiology*, 4th ed., 1988; International Commission on Microbiological Specifications for Foods, *Microbial Ecology of Foods*, vol. 1: *Factors Affecting Life and Death of Microorganisms*, 1980; J. M. Jay, *Modern Food Microbiology*, 6th ed., 2000; T. J. Montville (ed.), *Food Microbiology,* vol. 2: *New and Emerging Technologies*, 1987; J. L. Oblinger (ed.), Bacteria associated with foodborne diseases, *Food Technol.*, 42(4):181–200, 1988; B. P. Wasserman, T. J. Montville, and E. L. Korwek, Food biotechnology, *Food Technol.,* 42(1):133–146, 1988.

Food poisoning

An acute gastrointestinal or neurologic disorder caused by bacteria or their toxic products, by viruses, or by harmful chemicals in foods.

Bacteria may produce food poisoning by three means: (1) they infect the individual following consumption of the contaminated food; (2) as they grow they produce a toxin in food before it is consumed; or (3) they produce toxin in the gastrointestinal tract after the individual consumes the contaminated food.

Infectious bacteria associated with food poisoning include *Brucella*, *Campylobacter jejuni*, enteroinvasive *Escherichia coli*, enterohemorrhagic *E. coli*, *Listeria monocytogenes*, *Salmonella*, *Shigella*, *Vibrio parahaemolyticus*, *Vibrio vulnificus*, and *Yersinia enterocolitica*. These organisms must be ingested for poisoning to occur, and in many instances only a few cells need be consumed to initiate a gastrointestinal infection. *Salmonella* and *C. jejuni* are the most prevalent causes of food-borne bacterial infections. *See* YERSINIA.

Staphylococcus aureus and *Clostridium botulinum* are bacteria responsible for food poisonings resulting from ingestion of preformed toxin. *Staphylococcus aureus* produces heat-stable toxins that remain active in foods after cooking. Staphylococcal poisoning accounts for 20–40% of food-borne outbreaks reported in the United States. *Clostridium botulinum* produces one of the most potent toxins known. Botulinal toxin causes neuromuscular paralysis, often resulting in respiratory failure and death. *See* BOTULISM; STAPHYLOCOCCUS; TOXIN.

Food-poisoning bacteria that produce toxin in the gastrointestinal tract following their ingestion include *Bacillus cereus*, *Clostridium perfringens*, enterotoxigenic *E. coli*, and *Vibrio cholerae*. *Bacillus cereus* and *C. perfringens* are spore-forming bacteria that often survive cooking and grow to large numbers in improperly refrigerated foods. Following ingestion, their cells release enterotoxins in the intestinal tract. Enterotoxigenic *E. coli* is a leading cause of travelers' diarrhea. *See* DIARRHEA; ESCHERICHIA.

Viruses that cause food-borne disease generally emanate from the human intestine and contaminate food through mishandling by an infected individual, or by way of water or sewage contaminated with human feces. Hepatitis A virus and Norwalk-like virus are the preeminent viruses associated with food-borne illness. *See* HEPATITIS.

Chemical-induced food poisoning is generally characterized by a rapid onset of symptoms which include nausea and vomiting. Foods contaminated with high levels of heavy metals, insecticides, or pesticides have caused illness following ingestion. *See* FOOD MICROBIOLOGY; MEDICAL BACTERIOLOGY; TOXICOLOGY. Mike Doyle

Bibliography. F. L. Bryan, *Diseases Transmitted by Foods*, 2d ed., Centers for Disease Control, Atlanta, 1982; M. P. Doyle, *Foodborne Bacterial Pathogens*, 1989; M. P. Doyle and E. M. Foster, Bacterial agents of foodborne disease, *Wis. Med. J.*, 80:24–27, 1981; Y. H. Hui, M. Gorham, and Cliver (eds.), *Foodborne Disease Handbook: Diseases Caused by Hazardous Substances*, 1994.

Food preservation

The branch of food science and technology that deals with the practical control of factors capable of adversely affecting the safety, nutritive value, appearance, texture, flavor, and keeping qualities of raw and processed foods. Since thousands of food products differing in physical, chemical, and biological properties can undergo deterioration from such diverse causes as microorganisms, natural food enzymes, insects and rodents, industrial contaminants, heat, cold, light, oxygen, moisture, dryness, and storage time, food preservation methods differ widely and are optimized for specific products. Apart from application of a single food preservation method, a combination of methods may be used to improve the safety and storage stability of foods. This concept is currently referred to as hurdle technology; it is also known as food preservation by combined methods, combination preservation, or combination techniques.

Food preservation methods involve the use of heat, refrigeration, freezing, concentration, dehydration, radiation, pH control, chemical preservatives, and packaging applied to produce various degrees of preservation in accordance with the differing use patterns and shelf-life needs of unique products.

Perishability of many food materials was somewhat controlled long before the principles of modern food preservation were understood. Cheese and other fermented milk products, wine, sauerkraut and pickles, smoked meats and fish, dried and sugared fruits, and numerous other foods had their beginnings in attempts to extend the storage life of the basic commodities from which they were derived, but results were often disappointing.

Optimum food preservation must eliminate or minimize all of the factors that may cause a given food to deteriorate, without producing undue adverse effects. This can be especially difficult since the components of foods may be more sensitive to preservation treatments than the highly resistant bacterial spores and natural food enzymes targeted for destruction. Many nonbiological causes of food deterioration must be prevented also; these include oxygen, light, and loss of moisture.

While traditional methods of food preservation, including heating, freezing, drying, refrigerating, and acidifying, are still widely used in the food industry, new food preservation methods are actively being researched. These include physical methods such as high hydrostatic pressure, pulsed electric fields, ohmic heating, and ultrasound. Such methods have been shown to be effective in preserving food with minimal effects on nutrients and sensory quality that might otherwise be destroyed by traditional food preservation methods such as heating, drying, and acidification.

Heat. Thermal processes to preserve foods vary in intensity. True sterility to ensure total destruction of the most heat-resistant bacterial spores in nonacidic foods may require a treatment of at least 250°F (121°C) of wet heat for at least 15 min, or its lethal equivalent, to be delivered throughout the entire food mass. Such a treatment would be damaging to most foods. The term commercial sterility refers to a less severe condition that still assures destruction of all pathogenic organisms, as well as organisms that, if present, could grow in the product and produce spoilage under normal conditions of handling and storage. Most of the canned food supply that is stable at room temperature is commercially sterile. This is commonly achieved in canning retorts with steam under pressure at temperatures and for times that vary, depending upon container size and chemical and physical properties of the food, which can affect heat-transfer rates and the thermal resistance of organisms. *See* STERILIZATION.

Many foods are subjected to still less severe heating by methods that produce pasteurization to assure destruction of pathogens and extend product shelf life by inactivating food enzymes and reducing the number of spoilage organisms. Pasteurization of milk is achieved with a temperature of 145°F (63°C) for 30 min, or its thermal lethal equivalent. Since significant numbers of nonpathogenic bacteria survive, storage life is extended by refrigerating the pasteurized milk. Beer, wine, fruit juices, and other foods are commonly pasteurized, but at different temperatures. Heat blanching is a kind of pasteurization applied to vegetables to inactivate enzymes when such products are to be frozen, since frozen storage of itself does not stop enzyme activity. *See* PASTEURIZATION.

The lethality of heat always depends upon temperature and time. Higher temperatures for shorter times can be as effective as lower temperatures for longer times, and appropriate combinations can be selected for thermal lethal equivalency. Time–temperature combinations with equivalent microbial lethality, however, are not equal with respect to the damaging effects these can have on color, flavor, texture, and nutritive values of foods. In this regard, higher temperatures for shorter times will yield products superior to those produced with lower temperatures for longer times.

Advances in thermal processing incorporate the high-temperature short-time (HTST) principle whether pasteurization or commercial sterilization is the goal. The application of high-temperature short-time processing is more easily accomplished with liquid foods or liquids containing small particulates than with solid foods, since the former can easily be heated and cooled rapidly by passing them in thin layers through specially designed heat exchangers. This is done in the process of aseptic canning, where products prepared to commercial sterility standards are heated to temperatures as high as 302°F (150°C) for 1 or 2 s and as quickly cooled, and then sealed in previously sterilized containers within an aseptic environment.

Cooling and freezing. The slowing of biological and chemical activity with decreasing temperature is the principle behind cooling (refrigeration) and freezing preservation. In addition, when water is converted to ice, free water required for its solvent properties by all living systems is removed. Even severe freezing, however, will not destroy large numbers of microorganisms or completely inactivate food enzymes; these can resume rapid activity, unless inhibited by other means, when food is removed from cold or frozen storage.

Most microorganisms grow best in the range of about 60–100°F (16–38°C). Psychrotrophic bacteria thrive at low temperatures and can grow slowly at temperatures down to 32°F (0°C) and below if free water exist. Most pathogens cannot grow below 40°F (4°C). Home refrigerators commonly operate in the range of about 40–45°F (4–7°C). Some fruits and vegetables store best at temperatures of about 50°F (10°C), and commercial refrigerated storage may be optimized for specific products. Refrigerated storage life of many foods can be extended by the use of packaging that minimizes moisture loss and controls gas atmospheres within packages.

Highest-quality frozen foods depend upon very fast rates of freezing. Slow freezing leads to the growth of large irregular ice crystals capable of disruption of delicate food textures. Slow freezing also increases the time during which food constituents can react adversely with solutes that become concentrated by liquid water changing to ice as freezing progresses. Thus, rapid freezing has been the goal of advanced freezing processes.

Commercial freezing methods utilize refrigerated still air; high-velocity air, which is faster and more efficient; and high-velocity air made to suspend particulate foods, such as peas, as in a fluidized-bed fast freezer. Indirect-contact freezing utilizes hollow flat plates chilled with an internally circulated refrigerant to freeze solid foods, or with refrigerated tubular heat exchangers that rapidly slush-freeze liquids. Immersion freezing involves direct contact of the food or its container with such refrigerants as cold brine, a glycol approved for food, or a fast-freezing cryogenic liquid, such as liquefied carbon dioxide or liquid nitrogen. Liquid nitrogen has a temperature of −320°F (−196°C). *See* COLD STORAGE.

Concentration and dehydration. When sufficient water is removed from foods, microorganisms will not grow, and many enzymatic and nonenzymatic reactions will cease or be markedly slowed. Free water that can enter into biological and chemical reactions is more important than total water, since some water may be bound and unavailable to support deteriorative processes. Free water exerts vapor pressure and possesses water activity (that is, provides water for bacterial growth), which must be decreased below critical levels if foods are to be preserved. Sugar syrups are concentrated foods whose water activity is below that required to support microbial spoilage. Sugar added to fruit juice will bind water, lower the juice's water activity, and, in sufficient concentration, yield a jelly that does not undergo microbial spoilage at room temperature. Concentration preservation, therefore, can be achieved by physically removing water, as by boiling or with lower-temperature vacuum evaporation, or by binding water through the addition of sugar, salt, or other solutes.

Foods preserved by dehydration contain considerably lower water activity and less total water than concentrated foods. Sun-dried cereal grains contain about 14% total water. Most dehydrated foods such as dried milk, instant coffee, and dehydrated potato flakes or granules contain less than 10% total water, and some, such as fruit juice crystals, contain less than 2%.

Most dehydration methods utilize heat to vaporize and remove water. This is most efficiently achieved when a food can be highly subdivided to produce a large surface area for rapid heat transfer into the food and rapid moisture transfer out. Liquid foods and purees commonly are atomized into a heated chamber (as in spray drying), spread thinly over the surface of a revolving heated drum from which they are continuously scraped (as in drum drying), and sometimes thickened or foamed and cast on belts that move through a tunnel oven. Solid foods may be diced to uniform piece size for more even drying and dried with heated moving air in cabinets, on belts, or within rotating cylinders to provide tumbling action. Fluidized-bed dryers use high-velocity air to suspend particulates for still faster drying.

The heat and oxygen sensitivity of many foods necessitates vacuum dehydration for high quality. Under vacuum, water can be removed at reduced temperature, and oxidative changes are minimized.

Solid foods tend to shrink and undergo shape distortion when they are dried. This can be overcome by freeze-drying whereby foods are frozen quickly and placed in a chamber under high vacuum. Vacuum and temperature conditions are regulated to promote sublimation of water vapor from the ice phase without the ice melting. The food's structure remains rigid as it goes directly from the frozen state to dryness. Because of its gentleness, freeze-drying is also used to dehydrate liquid foods such as coffee. A disadvantage of freeze-drying, however, is that it is more costly than other drying methods. *See* DRYING; SUBLIMATION.

Irradiation. X-rays, ultraviolet light, and ionizing radiations (including gamma and beta rays) belong to the electromagnetic spectrum of radiations and differ in frequency, wavelength, penetrating power, and the effects upon biological and nonbiological systems. Ionizing radiations may be obtained from radioactive isotopes, such as cobalt-60, or from electron accelerator machines. These radiations penetrate foods and exert their major effects by producing free radicals from water and other substrates. Depending upon dose intensity, these radiations can inhibit the sprouting of tubers, destroy insects, inactive some enzymes, and kill microorganisms to the point of pasteurization or sterilization. Preservation of food using sterilizing irradiation doses is done for various purposes, such as for use by astronauts during space missions and for immune-compromised persons who are highly susceptible to microbial infection. *See* FREE RADICAL.

Food irradiation remains highly controversial, partly because of fears that the safety of products and processes cannot be adequately regulated. In the United States, treatment of spices, raw or frozen beef, pork, or poultry meat to destroy microbial contamination is among the very few applications that are permitted. Several other countries permit wider use of food irradiation, including low-dose irradiation pasteurization to extend the storage life of highly perishable fruits and vegetables, poultry, and seafoods.

pH control. Hydrogen-ion concentration affects the rate and course of a great variety of chemical reactions. Microbial growth and metabolism and the activities of food enzymes exhibit pH optima and can be controlled to various degrees beyond these optima. The natural acids of certain fruits and vegetables, acid added as a chemical, and acid produced by fermentation can inhibit or partially inhibit several pathogenic and spoilage organisms. *Clostridium botulinum*, the most heat-resistant pathogen found in foods, will not grow and produce toxin at a pH of 4.6 or below. Therefore, foods with a pH in this range do not constitute a health hazard from this organism, and they do not require heat processing as severe as that required for more alkaline foods. Further, acid enhances the lethality of heat, often permitting milder heating conditions.

The pH of acidic foods, however, is rarely sufficiently low to assure long-term preservation from acid alone. Many acidic and fermented foods further depend upon prior pasteurization of their ingredients, the addition of salt and other chemicals, and refrigeration. *See* PH.

Chemical preservative. The U.S. Food and Drug Administration and comparable agencies in various countries vigorously regulate the chemicals that may be added to foods as well as the conditions of their use. Chemical preservatives and similar substances include antimicrobials, such as sodium benzoate, sorbic acid, and sodium nitrite; enzyme inhibitors, such as sulfur dioxide, to control browning of fruits and vegetables; and antioxidants, including butylated hydroxyanisole (BHA) and butylated hydroxytoluene (BHT), to help control fat rancidity. New chemicals must undergo rigorous testing to be approved, and approved chemicals may subsequently be prohibited when new information relative to safety warrants such action. There is much pressure to remove chemicals from the food supply, especially where their effects can be achieved by other means.

Packaging. Preservation methods cannot be effective without adequate packaging. Packaging protects foods from contamination, moisture gain or loss, flavor loss and odor pickup, the adverse effects of light, physical damage, and intentional tampering. Packaging enables food to be stored under vacuum, inert gases, or carefully selected gases that can control respiration of fruits and vegetables, biological changes in meat, and growth of microorganisms. Packages and packing materials must be carefully chosen to withstand the stresses of heating, freezing, and other operations since many products are processed within their final containers. Ultimately, a food product's quality and storage life are determined largely by its package. *See* FOOD ENGINEERING; FOOD MANUFACTURING; FOOD MICROBIOLOGY.

Combined preservation methods. The application of combined food preservation methods is called hurdle technology. This approach to food preservation is based on the use of a combination of several preservative factors (called hurdles) such as high or low temperature, reduced water activity, acidity, and antimicrobial food additives. These hurdles inhibit or prevent growth of food-borne microorganisms to enhance food safety and reduce food spoilage. The application of hurdle technology permits mild but very effective preservation of safe, stable, nutritious foods with desirable sensory quality because only a relatively low level of each preservation factor is usually required to improve the storage stability of foods.

Aubrey F. Mendonca; Norman N. Potter

Bibliography. G. V. Barbosa-Canovas et al., *Nonthermal Preservation of Foods*, Marcel Dekker, New York, 1998; L. Leistner and G. Gould, *Hurdle Technologies: Combination Treatments for Food Stability, Safety and Quality*, Kluwer Academic/Plenum, New York, 2002; D. G. Olson, Irradiation of food, *Food Technol.*, 52:56–62, 1998; N. N. Potter and J. H. Hotchkiss, *Food Science*, 5th ed., Aspen Publishers, Gaithersburg, 1998; M. Shafiur Rahman (ed.), *Handbook of Food Preservation*, Marcel Dekker, New York, 1999; P. Zeuthen and L. Bogh-Sorensen, *Food Preservation Techniques*, CRC Press, Boca Raton, 2003.

Food science

A science dealing with the physical, chemical, and biological properties of foods, in addition to the factors affecting them and their ultimate effects upon the sensory, nutritional, and storage properties and the safety of foods. Food properties are influenced by growing, harvesting, and slaughtering practices, preservation and preparation methods, processing and storage conditions, and packaging. Food science and its applications must be further concerned with economics and marketing; food preferences of various populations; quality assurance and control; regulatory aspects dealing with safety, wholesomeness, and honest representation; and the production of affordable, quality food on a worldwide basis. Therefore, food science interfaces with and draws upon many disciplines, including chemistry, physics, mathematics, the plant and animal sciences, biochemistry, enzymology, microbiology, genetics, engineering, statistics, computer science, nutrition, toxicology, psychology, and law.

Food science is concerned with food composition, which includes not only the kinds and amounts of food constituents (such as proteins, fats, carbohydrates, vitamins, minerals, emulsifiers, acids, oxidants, antioxidants, enzymes, pigments, flavors, and water) but also their physical and physicochemical arrangements. These constituents determine the properties of color, texture, flavor, nutritive value, and keeping quality of food. Each property may be grossly changed by processing. That is, processing may alter a food's cellular integrity and the states of colloidal dispersion, emulsification, or solution of its constituents (as well as its constituent interactions), and the susceptibility to deterioration by air, light, temperature, relative humidity, microorganisms, and natural food enzymes.

Food science deals with many food commodities and thousands of derived products. These commodities are processed for many reasons, including preservation, creation of new product forms, improvement of sensory and nutritional qualities, convenience, and removal of natural toxicants. Common processes are heating, cooling, freezing, concentration, dehydration, fermentation, sometimes irradiation, and packaging. Each process can be further classified by method; dehydration, for example, may employ drum drying, spray drying, cabinet drying, tunnel drying, fluidized-bed drying, vacuum drying, foam drying, freeze drying, and osmotic drying. Each process utilizes unique equipment that requires optimization of variables to maximize product quality and minimize costs, and is capable of yielding product with distinct characteristics. Further, commodities are increasingly utilized for their constituents which are separated, extracted, chemically and physically modified, and then recombined into an endless variety of formulated and engineered foods that are indistinguishable from their source materials. This often requires the use of highly specific food additives to improve processing properties and acceptance factors, including nutritional quality. Nutrient levels and nutrient availability from foods may be decreased or increased by handling and processing practices.

Food science is also concerned with all aspects of food safety, including natural food toxicants, industrial contaminants, misuse of food additives, and microbiological contamination, as well as methods for the detection, exclusion, inactivation, removal, and regulation of harmful substances. Consumers also must be protected against economic deception and misrepresentation.

Undergraduate and graduate degrees in food science are offered by universities around the world. Food scientists, food technologists, and professionals from related fields are employed throughout the food industry and by suppliers of equipment, ingredients, packaging materials, and other services, as well as by government and academia. Food scientists belong to many professional organizations; predominant among them is the Institute of Food Technologists. This organization, whose membership is international, is an excellent resource for information on all aspects of food science. *See* FOOD ENGINEERING; FOOD MANUFACTURING. **Norman N. Potter**

Bibliography. A. H. Ensminger et al., *Foods and Nutrition Encyclopedia*, 2d ed., 1993; Y. H. Hui, *Encyclopedia of Food Science and Technology*, 4 vols., 1991; N. N. Potter, *Food Science*, 5th ed., 1998.

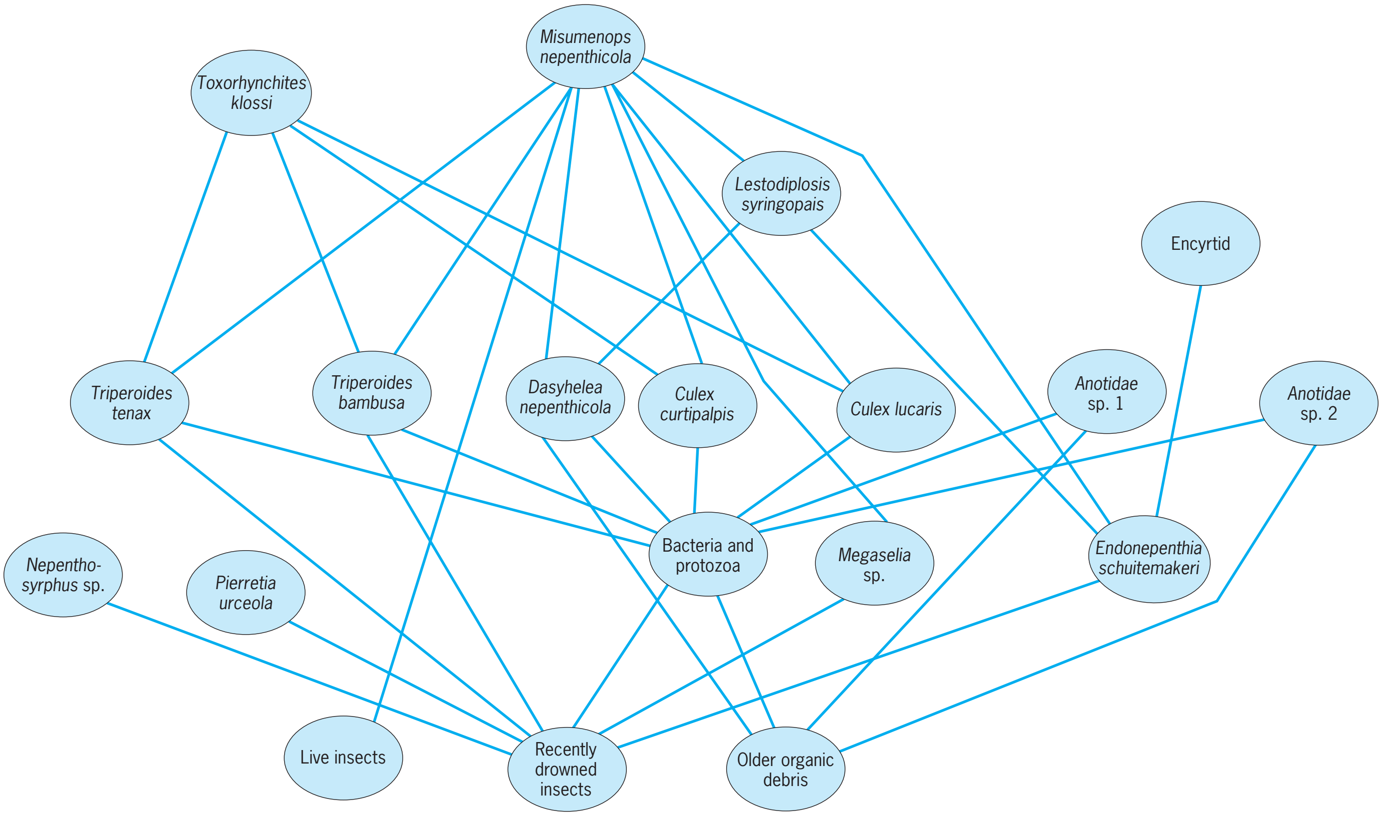
Food web

A diagram depicting those organisms that eat other organisms in the same ecosystem. In some cases, the organisms may already be dead. Thus, a food web is a network of energy flows in and out of the ecosystem of interest. Such flows can be very large, and some ecosystems depend almost entirely on energy that is imported. A food chain is one particular route through a food web.

A food web helps depict how an ecosystem is structured and functions. Most published food webs omit predation on minor species, the quantities of food consumed, the temporal variation of the flows, and many other details.

Example. Pitcher plants (belonging to the family Nepenthaceae) are tropical species with a structure that holds water into which insects fall and drown. The plant gains nutrients from the decaying insects. Many other species thrive in these tiny ponds, grabbing the energy from the decaying insects before the plant can do so (see **illus.**).

At the base of the food web in the pitcher plant are drowning insects, dead insects, and older organic debris. Some of the species that feed on the decaying material have predators, others do not. In addition, some species may receive energy from more than one level. One predator, a spider (genus *Misumenops*), feeds on another predator, a fly (*Lestodiplosis*), and both these species feed on another fly (*Endonepenthia*). Thus, species are not always arranged into neat, clear food chains.



Food web of the insects found in the pitcher plant, *Nepenthes albomarginata*, in West Malaysia. Each line represents a trophic linkage; predators are higher in the web than their prey.

Food chains. Along a simple food chain, A eats B, B eats C, and so on. For example, the energy that plants capture from the sun during photosynthesis may end up in the tissues of a hawk. It gets there via a bird that the hawk has eaten, the insects that were eaten by the bird, and the plants on which the insects fed. Each stage of the food chain is called a trophic level. More generally, the trophic levels are separated into producers (the plants), herbivores or primary consumers (the insects), carnivores or secondary consumers (the bird), and top carnivores or tertiary consumers (the hawk).

Food chains may involve parasites as well as predators. The lice feeding in the feathers of the hawk are yet another trophic level. When decaying vegetation, dead animals, or both are the energy sources, the food chains are described as detrital.

Patterns. Food webs have a long history in ecology, including C. Darwin's famous description of a "tangled bank" at the end of *Origin of Species*. This analogy to tangles implies that even when food webs ignore some details, they demonstrate how complex nature can be. This complexity raises questions: First, are food web patterns simply random, or is there a pattern to their tangles? Second, what are the consequences of these patterns for the dynamics of the ecosystem?

Food chains are usually short. The shortest food chains have two levels. For example, in the illustration there are the drowning insects in the water, and the spider *Misumenops*. The longest chains have five levels [old organic debris, the bacteria and protozoa that feed on it, two species of *Culex* (mosquitoes), *Dasyhelea* (biting midges), and *Misumenops*]. One way to describe and simplify various food chains is to count the most common number of levels from the top to the bottom of the web. Thus, although *Misumenops* sits atop chains of length two and five, the most common chains in the web are those of length three, which is quite typical. Most food chains are three or four trophic levels long (if parasites are excluded), though there are longer ones.

Energetics. There are several possible explanations for why food chains are generally short. The first involves energy. In many ecosystems there are more plants than insects, more insects than insectivorous birds, and more insectivorous birds than hawks. These ecological pyramids reflect an underlying energetic constraint. The first law of thermodynamics states that when energy is converted from one form to another the amount of energy remains constant. An approximate statement of the second law is that the amount of useful energy decreases at each conversion. When insects eat plants, they convert energy locked in plant tissues into insect tissue. Yet, they pay an energetic cost in doing so. Mammals and birds pay an even greater cost. Only about 1% of the energy of the food consumed goes to produce new tissues. The rest is lost as heat, both for warmth and as a by-product of the conversion process. *See* ECOLOGICAL ENERGETICS.

Thus, between each trophic level, much of the energy is lost as heat. As the energy passes up the food chain, there is less and less to go around. There may not be enough energy to support a viable population of a species at trophic level five or higher.

This energy flow hypothesis is widely supported, but it is also criticized because it predicts that food chains should be shorter in energetically poor ecosystems such as a bleak arctic tundra or extreme deserts. These systems often have food chains similar in length to energetically more productive systems.

Recovery from disaster. Another hypothesis about the shortness of food chains has to do with how quickly particular species recover from environmental disasters. For example, in a lake with phytoplankton, zooplankton, and fish, when the phytoplankton decline the zooplankton will also decline, followed by the fish. The phytoplankton may recover but will remain at low levels, kept there by the zooplankton. At least transiently, the zooplankton may reach higher than normal levels because the fish, their predators, are still scarce. The phytoplankton will not completely recover until all the species in the food chain have recovered. Mathematical models can expand such arguments. These models show that the longer a food chain, the longer it will take its constituent species to recover from perturbations. (The phytoplankton could recover quickly in the example if they were the only trophic level.) Species atop very long food chains may not recover before the next disaster. Such arguments predict that food chains will be longer when environmental disasters are rare, short when they are common, and will not necessarily be related to the amount of energy entering the system.

Consequences for species dynamics. The number of trophic levels a food web contains will determine what happens when an ecosystem is subjected to a short, sharp shock—for example, when a large number of individuals of one species are killed by a natural disaster or an incident of human-made pollution—and how quickly the system will recover.

The food web will also influence what happens if the abundance of a species is permanently reduced (perhaps because of harvesting) or increased (perhaps by increasing an essential nutrient for a plant). For example, the population of a fertilized plant species may not expand because herbivores consume the increased plant production, and then the more numerous herbivores may be consumed by their predators, and so on up the food chain.

Some species have redundant roles in an ecosystem so that their loss will not seriously impair the system's dynamics. Therefore, the loss of such species from an ecosystem will not have a substantial effect on ecosystem function. The alternative hypothesis is that more diverse ecosystems could have a greater chance of containing species that survive or that can even thrive during a disturbance that kills off other species. Highly connected and simple food webs differ in their responses to disturbances, so once again the structure of food webs makes a difference. *See* ECOLOGICAL COMMUNITIES; ECOSYSTEM; POPULATION ECOLOGY.

Stuart Pimm

Bibliography. S. R. Carpenter, *Complex Interactions in Lake Communities*, 1988; J. E. Cohen, F. Briand, and C. M. Newman, *Community Food Webs*, 1990; S. L. Pimm, *Food Webs*, 1982; S. L. Pimm, J. H. Lawton, and J. E. Cohen, Food web patterns and their consequences, *Nature*, 350:669–674, 1991.

Foot-and-mouth disease

A highly contagious viral disease of domesticated and wild cloven-hoofed animals with the potential to cause enormous economic losses. It is characterized by the formation of vesicles on the feet, in and around the mouth (**Fig. 1**), and on the mammary gland. At the acute stage there is high fever, depression, lameness, and reduced appetite. Milking animals show a sudden reduction in production. The mortality in adult animals is usually less than 3%, but in young animals it can exceed 90%.

Properties of virus. The causative virus is a member of the *Aphthovirus* genus of the family Picornaviridae. The virion ranges 20 to 30 nanometers in diameter and contains a positive-stranded ribonucleic acid (RNA) genome. The nonenveloped viral capsid is icosahedral in shape and consists of 60 copies of each of the four structural proteins VP1–4 (**Fig. 2**). The virus is acid-labile and inactivated outside the pH range 6.0–8.0.

Seven serotypes are recognized: O, A, C, SAT 1, SAT 2, SAT 3, and Asia I. There is no cross-immunity between serotypes. Each serotype contains strains of varying antigenic and genomic characteristics.

Pathogenesis. The virus can infect by different routes and mechanisms. Ruminant animals are usually infected through the upper respiratory tract following the inhalation of virus associated with droplets or aerosol particles. Pigs are most often infected through the alimentary tract following the ingestion of virus in contaminated foodstuffs (such as meat or milk). In cattle infected by the respiratory route, the virus replicates initially in the mucosae and lymphoid tissues of the pharyngeal region and associated lymph nodes. It then enters the bloodstream and infects glandular organs and lymph nodes

Fig. 1. Mouth of a steer with foot-and-mouth disease lesions on the tongue, along the lower gum, and inside the lower lip.

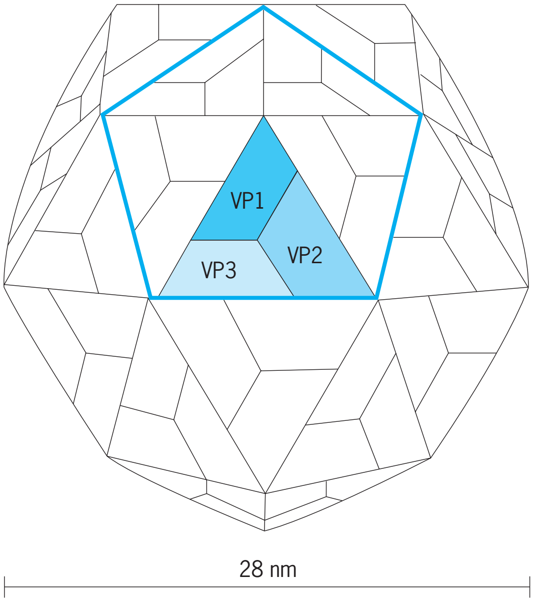


Fig. 2. Schematic diagram of the capsid of a foot-and-mouth disease virion. The four structural proteins, VP1-4 (also called 1A, 1B, 1C, and 1D), make up a protein subunit or protomer. Five protomers join to form a pentamer, and a capsid is an assembly of 12 protomers. VP4 (1A) is located internally beneath the other three proteins.

throughout the body. Other predilection sites, such as the epithelial tissues of the mouth, feet, and mammary gland, are infected during this phase, leading to vesicle formation. The virus has a high affinity for the cardiac tissue of young animals, in which it can cause a severe pathology leading to heart failure and death.

During the acute phase of disease, large amounts of virus occur throughout the tissues and organs of the infected animal and in its excretions and secretions. After 4–5 days, circulating antibody appears, and there is rapid clearance of virus from all sites and fluids. An exception is the pharyngeal region of ruminants, where virus can persist in a proportion of animals for months or years. The duration of infection is species-related; the African buffalo can remain infected for 5 years, cattle for $3^1/_2$ years, and sheep for 9 months.

Epidemiology. The movement of infected animals and the transmission of virus by contact to susceptible animals in a new herd or flock is by far the most common mechanism of spread. Next in frequency is the movement of contaminated animal products, such as meat and milk—most likely to infect pigs through ingestion. Once a few pigs have been infected, the quantity of virus will be greatly amplified, and spread by direct contact will be possible. More rarely, spread results from the mechanical transfer of virus by people, by vehicles, or on fomites (any article that has been in contact with a virus and is capable of transmitting it). Another infrequent mechanism is dispersal of the virus by wind. This generally requires a series of special circumstances: infected pigs at source, a flat terrain, a low-to-moderate wind speed, low atmospheric turbulence, a high humidity, and a high density of cattle downwind. The most dramatic instances of airborne spread of foot-and-mouth disease have been recorded in northwestern Europe, where windborne spread is believed to have taken place over distances of 60 km (37 mi) over land and 300 km (186 mi) across the sea.

Immune response. Immunity to the disease is primarily mediated through neutralizing antibodies directed against the structural proteins of the virus, especially VP1. Antigen-specific T cells play an essential part in the induction and regulation of the circulating antibody response. Specific antibodies are detectable by bioassay 4–5 days after infection, at which time viremia ceases and there is a progressive reduction in virus excretion as the lesions heal. Recovered animals are immune to challenge by the homologous strain of virus for over 4 years.

Laboratory diagnosis. The detection of the viral antigen is sufficient for a positive diagnosis. The specimen of choice is epithelial tissue from an unruptured or freshly ruptured vesicle. Normally, viral antigen is detected and serotyped by enzyme-linked immunosorbent assay (ELISA). In cases where it is not possible to collect epithelial tissue, the demonstration of an increased titre in sequentially collected paired serum samples from nonvaccinated animals is sufficient for a positive diagnosis. Virus neutralization and ELISA are used for serological tests. *See* SEROLOGY.

Control. Countries free from foot-and-mouth disease take stringent measures to prevent the entry of virus. Normally they import animals and animal products only from other countries free of the disease. If they decide to import from a country of lower health status, strict safeguards are imposed to manage the risk. When an outbreak occurs in a country free of the disease, the usual procedure is to apply the stamping-out policy. This involves the slaughter of all affected and in-contact cloven-hoofed animals on the infected premises and the disposal of the carcasses by burning or burial. This step is followed by a thorough cleansing and disinfection of the premises. A movement standstill is enforced in the surrounding area until it is evident that virus is no longer circulating.

Control in countries where the disease is endemic or sporadic is mainly by vaccination. The vaccines contain inactivated viral antigens of one or more serotypes, depending on the prevailing disease situation. Immunity following primary and booster doses of vaccine lasts for around 6 months. *See* ANIMAL VIRUS; EPIDEMIOLOGY.

Alex I. Donaldson

Bibliography. G. R. Thomson, Foot-and-mouth disease, in J. A. W. Coetzer, G. R. Thomson, and R. C. Tustin (eds.), *Infectious Diseases of Livestock with Special Reference to Southern Africa*, vol. II, Oxford University Press, 1994.

Foot disorders

Musculoskeletal, neurological, or dermatological abnormalities of the foot. They may be of developmental or acquired origin.

Disorders of bones and joints. Musculoskeletal foot problems may be congenital or secondary to neurological disorders, joint laxity, infection, arthritis, or tendinopathy (tendon injury). Poorly fitting footwear may exacerbate symptoms from a deformed foot but is not in itself a cause of foot deformity. *See* JOINT (ANATOMY); TENDON.

Clubfoot. Clubfoot is a congenital deformity that is also known as talipes. Talipes calcaneovalgus, in which the sole of the foot is turned upward and outward, is usually a mobile deformity and has a good prognosis even if untreated. Talipes equinovarus, in which the sole of the foot is turned downward and inward, requires early treatment, especially when the deformity is rigid. The incidence in the United States is 1 in 1000 live births but is lower in some countries such as China and much higher in other areas, especially Polynesia. Almost half of cases are bilateral.

If untreated the deformity persists so that the child has to walk on the top of the foot. The results of early manipulation and use of a serial plaster cast are good. If this fails to give good correction, surgical release of the contracted tendons and ligaments is necessary to prevent the development of secondary bony deformity. *See* CONGENITAL ANOMALIES.

Flatfoot. Flatfoot, or pronated foot, is a condition in which the longitudinal arch of the foot is lost. Although it gives rise to little disability in childhood, it is a common cause of maternal concern, and it used to be a reason for rejection from military service. In this condition, which is not seen until a child starts to walk, the heel of the foot is seen to go into a valgus position (turned outward) and this corrects well when the child stands on tiptoe. The deformity is usually hypermobile and associated with excessive generalized joint laxity. However, fixed flat foot deformity may have more serious underlying causes such as congenital coalition of the hindfoot bones.

The prognosis for flatfoot is good. Most cases are treated by an insole (orthosis), and surgery is rarely indicated.

A suddenly occurring flatfoot deformity in the adult may occur in middle-aged females and is usually due to tendinopathy or rupture of the tibialis posterior tendon on the inside of the foot and ankle. If the tendon is ruptured, early surgical treatment is essential.

Bunion. A bunion (hallux valgus) is a deformity in which the big toe is deviated away from the midline, and the first metatarsal bone moves away from adjacent second metatarsal with splaying out of the forefoot. The ensuing prominence of the first metatarsal head at the base of the big toe gives rise to the apparent bony deformity or bunion. Wearing tight, narrow, high-heeled shoes makes this a painful condition in many middle-aged females. Pressure of the shoe against this prominent bone may give rise to a protective sac of fluid (adventitious bursa). There may be an associated pronation of the foot, and treatment with an orthosis may be advised in mild cases.

Surgical treatment is indicated in cases with persistent pain, especially if the condition does not respond to modifications in footwear. Simple excision of the bunion is not usually successful, and a variety of surgical procedures to realign the first metatarsal bone and correct the position of the joint at the base of the big toe are used.

Hammertoe. This clawing of the toes may present in small children and is usually thought to be associated with a mild neuromuscular imbalance. Severe cases may occur in children with spina bifida, poliomyelitis, and Charcot-Marie-Tooth disease. Fixed deformity may develop, and in severe cases surgical treatment by release or transfer of tendons becomes necessary.

In the adult, hammertoe deformity may develop in the smaller toes without any obvious underlying neurological abnormality. In this deformity, the proximal phalanx of the toe (the one nearest to the rest of the foot) becomes extended upward, the middle phalanx flexed down, and the distal phalanx (the one farthest from the rest of the foot) hyperextends to remain parallel with the floor. Pain is due to pressure on the joint between the proximal and distal phalanges and may be relieved by fitting a large toe box. In severe cases, surgical treatment to fuse this joint (arthrodesis) may be advised.

Heel spur and plantar fasciitis. The plantar fascia is a strong ligament that maintains the longitudinal arch of the foot and acts like a bow string with attachment to the heel (calcaneus) and forefoot. It may become inflamed at the calcaneal attachment, especially in athletes and overweight patients, when excessive forces are applied to this area. In some cases, there is an outgrowth of bone in this area (calcaneal spur), although a calcaneal spur may be an incidental finding on x-ray in someone without pain.

Treatment may involve a variety of measures, including stretching exercises for the Achilles tendon, use of an orthosis, local steroid injection, and extracorporeal shock wave therapy. If these measures fail, surgery to release the attachment of the plantar fascia may be indicated.

Calcaneal apophysitis. This condition, also known as Sever's disease, causes heel pain from the apophysis (or posterior center of ossification of the calcaneus) in children, and may represent a stress fracture of the calcaneus. It may occur as a result of excessive athletic activity. Most cases settle with rest and use of an orthosis. Immobilization in a plaster cast may also be effective.

Neurological disorders. Neurological disorders may affect the foot by causing muscle imbalance and sensory loss. For example, an increasingly common problem seen in diabetic patients is peripheral neuropathy (damage to the peripheral nervous system resulting in the loss of sensation) causing deformity due to neuropathic arthropathy (degeneration of the stress-bearing portion of a joint) and painless ulceration of the sole of the foot. Morton's neuroma and tarsal tunnel syndrome are examples of entrapment neuropathy in the foot. *See* DIABETES; NERVE; NERVOUS SYSTEM (VERTEBRATE).

Morton's neuroma. A neuroma is a benign tumor of nerve cells. Morton's neuroma, most common in female patients between the ages of 50 and 60, is characterized by the thickening of the third intermetatarsal nerve over a period of years, causing a small benign fusiform (spindle-shaped) tumor to eventually form, usually in the space between the third and fourth toes. Pain is felt in the space and may be associated with pain and pins and tingling radiating into the third and fourth toes (or second and third toes if the neuroma lies between the second and third metatarsal heads). A fusiform tumor may be diagnosed by an ultrasound scan.

Conservative treatment is advised and involves the use of an orthosis in the shoe and a local steroid injection in the space between the metatarsal heads. Entrapment of the nerve may be relieved surgically by dividing the intermetatarsal ligaments to perform decompression. More commonly, the benign tumor is excised, although there is a risk of recurrent symptoms due to the risk of a traumatic neuroma developing in the stump of the nerve.

Tarsal tunnel syndrome. The posterior tibial nerve runs into the sole of the foot from behind the inner aspect of the ankle joint and into a channel known as the tarsal tunnel. Entrapment of the nerve within this tunnel gives rise to tarsal tunnel syndrome, characterized by pain, and tingling and numbness radiating into the sole of the foot along the cutaneous distribution of the nerve. In some cases, entrapment is due to a bony swelling near the inside of the ankle or to a cystic swelling such as a ganglion.

The initial treatment consists of placement of an orthosis inside the shoe or local steroid injection into the tarsal tunnel. Measurement of conduction velocity in the posterior tibial nerve may be used to confirm the diagnosis; if the changes are severe, surgical treatment may be necessary to decompress the nerve.

Dermatological disorders. The skin of the foot undergoes stress from pressure and friction. In response, it protects itself with defensive modifications that themselves cause pain.

Corns and calluses. A corn, also known as a heloma, is a small, sharply circumscribed, conically shaped and deep-seated area of skin. Calluses are similar lesions but are flatter and more evenly distributed. Both are likely to occur in the sole of the foot in response to excessive pressure and consist largely of the protein keratin. This pressure is usually the result of a bony deformity, for example, a metatarsal head prolapsed into the sole of the foot giving rise to a painful callosity beneath the fibular condyle of the metatarsal head known as intractable plantar keratosis. An adventitious bursa may also form in response to excessive pressure.

Shaving the corn or callus does not address the underlying problem. Orthotic treatment should be advised to reduce localized pressure. If this fails, surgery to correct the underlying bony deformity may become necessary.

Warts. Warts are benign lesions of the skin induced by the human papilloma virus and are secondary in prevalence to corns and callosities. They can occur elsewhere in the body but are most common in the sole of the foot (plantar warts). Transmission from person to person may occur but depends on individual immunity. The virus lies in the epidermis and stimulates this superficial layer of the skin to produce excessive keratin. This causes pressure, further stimulating keratin production and irritating pain receptors in the underlying dermis.

Treatment involves applying agents such as salicylic acid to the skin with an occlusive dressing. Surgical removal may be effective but runs the risk of painful scarring.

Ingrown toenail (onychocryptosis). This penetration of the edge of the nail into the surrounding soft tissues may be due to hyperplasia (increase in size) of the nail groove or deformity of the nail plate; the latter may be associated with an osseous (bony) malformation of the big toe or a fungal infection. This condition is most often seen in teenage boys and may respond to improved attention to hygiene and correct cutting of the toe nail. Initially, there is pain, redness, and swelling on either or both borders of the nail. If chronic infection supervenes, removal of all or part of the nail may be necessary. The nail bed may be treated by phenol or by surgical excision to prevent further growth of the ingrowing border and subsequent recurrence. Radical procedures to perform a terminal amputation of the big toe should not be necessary. T. W. D. Smith

Forage crops

Grasses and legumes that make up grasslands and are used as forages for livestock. The grasslands represent an ancient renewable natural resource. They form 25% of the world's vegetation and occupy the largest area of any single plant type. They benefit humanity indirectly by providing food for both wild animals and domesticated livestock, some of which are ruminants that, because their digestive systems contain microorganisms, are able to digest fibrous forage material. Thus, the prime value of grassland areas lies in the meat, milk, or work produced by the livestock that graze on them. *See* GRASSLAND ECOSYSTEM.

Grasslands have other attributes as well. In order to withstand the tug and pull of the grazing animal, a forage plant must have an extensive root system, and this makes a contribution to soil fertility. When the plant is grazed or cut, the photosynthetic area that remains is not large enough to provide sugars to maintain root respiration, and so part of the root system dies and adds to the organic material in the soil and to soil structure. Such is the basis of many highly productive crop rotations that maintain soil fertility without expensive fertilizer applications. Most forage plants are long-lived perennials that can be defoliated repeatedly during a growing season. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Forages also protect the soil from erosion by both wind and water. In fact, where row crop farming has led to soil deterioration, as it did when dust bowl conditions prevailed in North America, forages are used to restore and stabilize the land. Many of grassland areas make an important contribution to water catchment. They stabilize land not only in high-rainfall zones but also on the banks of rivers and canals that bring water to areas of high population. Moreover, the green and pleasant areas covered by forage have an esthetic appeal to people. Grasses are used to beautify parks, airfields, playgrounds, and highways. *See* EROSION; LAWN AND TURF GRASSES.

Plant groups. The grasses and the legumes, both of which are used as forages, are divided into cool-season or temperate species and warm-season or tropical plants.

Grasses. The grasses are very numerous, herbaceous, monocotyledonous plants. They are the result of over 30 million years of coevolution between plants and the animals that graze upon them. During this time, certain mouthparts of animals were modified to permit grazing on different plant types. For example, sheep are able to graze very close to the ground, while cattle can use only plant growth over 2–3 in. (5-7.5 cm) tall. Grasses have developed low growing points, well below the bite level of grazing animals, so that the plant can regrow after defoliation. The plant's inflorescence is extruded rapidly and is vulnerable to the animal for only a minimum time. Grass seeds have awns, spikes, and hooks. Many are disseminated by becoming attached to an animal's coat. At maturity, some grass heads are unattractive to grazing animals, and so the seed is not eaten. When the seed is eaten, it may be capable of passing through the animal's digestive system without damage. The cool-season grasses are lower-yielding but provide forage material of higher quality than do the warm-season type. *See* GRASS CROPS.

Some commonly grown grasses are the following:

- Cool-season grasses
  - Bentgrasses (*Agrostis* spp.)
  - Bluegrasses (*Poa* spp.)
  - Smooth bromegrass (*Bromis inermis*)
  - Fescues (*Festuca* spp.)
  - Meadow foxtail (*Alopecurus pratensis*)
  - Orchardgrass (*Dactylis glomerata*)
  - Ryegrasses (*Lolium* spp.)
  - Wildryes (*Elymus* spp.)
  - Timothy (*Phleum pratense*)
  - Wheatgrasses (*Agropyron* spp.)
- Warm-season grasses
  - Bahiagrass (*Paspalum notatum*)
  - Bermudagrass (*Cynodon dactylon*)
  - Dallisgrass (*Paspalum dilatatum*)
  - Johnsongrass (*Sorghum halepense*)
  - Rhodegrass (*Chloris gayana*)

Legumes. The legumes have been used by farmers for over 6000 years. The reason they are so valuable to humanity is that their root systems are capable of forming a symbiotic relationship with bacteria of the genus *Rhizobium*. The plant provides the bacteria with sugars for growth and metabolism. The bacteria are able to fix atmospheric nitrogen and make it available to the plant in the form of nitrates. The plant builds these salts first into amino acids and then into proteins. Thus, a nitrogen atom from the air has passed along a short food chain from bacteria to plant and on to the animal. *See* LEGUME FORAGES; NITROGEN FIXATION.

The most commonly grown forage legume is alfalfa, which is used in both the temperate and subtropical areas. The warm-season legumes are important since, grown in a mixture with warm-season grasses, they improve forage quality. *See* ALFALFA.

Some of the legumes are the following:

- Temperate-season legumes
  - Alfalfa (*Medicago* spp.)
  - Birdsfoot trefoil (*Lotus corniculatus*)
  - True clovers (*Trifolium* spp.)
  - Sainfoin (*Onobrychis viciaefolia*)
  - Sweetclover (*Melilotus* spp.)
- Warm-season legumes
  - Crimson clover (*Trifolium nicaenatum*)
  - Kudzu (*Pueraria lobata*)
  - Subterranean clover (*Trifolium subteraneum*)

Forage quality. There are three ways in which forage quality may be measured. Farmers use a number of "visual" indices like color and odor which indicate that the hay or silage was well made and is still in good condition. For a pasture that is to be grazed, the quality criteria are the species present and their stage of maturity, for as plants age quality declines. The second indices are obtained from chemical analysis of forage material. These attempt to determine the undigestible part of the forage, usually cellulose fiber coated with lignin. This fraction is called the acid detergent fiber (ADF). A chemical analysis can also be used to determine elemental nitrogen. This is usually expressed as crude protein (CP), which contains several nonprotein compounds as well as protein (see **table**). The third types of forage quality indices are the most reliable and the most expensive to determine. These are obtained from the performance of the animals themselves, but unfortunately this is very individualized. The size, age, species, and the individual likes and dislikes of the animal will influence intake and production. Thus, large animal populations (10–20 animals) have to be fed over several months to determine animal intake and live weight gain. Intake contributes 70% to live weight gain, while digestibility accounts for 30%. The important component of digestibility is the rate of passage through the digestive system. Good-quality forages are rapidly digested, and the animal will then start to eat again, thereby increasing intake.

Establishment. All forage seeds are very small and so, in contrast to large-seeded row crops like corn

Crude protein (CP) and acid detergent fiber (ADF) values for common forages

Forage type	CP,%	ADF,%
Alfalfa hay	17.4	38.1
Timothy hay	8.1	37.1
Barley silage	9.7	33.8
Oats silage	8.4	38.0

or wheat, need a carefully prepared seedbed. For forages, the seedbed should be firm and have a fine soil structure. Firm soil ensures that neither root nor shoot grow through air pockets since, having few food reserves, the small seedling might not survive. A fine structure ensures good contact between soil and seed and facilitates water uptake as well as water movement in the soil. Given these conditions, the seed should be placed at a very shallow depth in the soil. Small seedlings are unable to reach the surface if planted deeply. The young forage plant has very few nutrient requirements during the first year, and fertilizers are banded, that is, they are placed in the soil directly below the seed. Forage seed establishment requires precision and specialized forage seeding equipment.

Forages crops are usually grown in species mixtures that contain a legume and one or two grasses. The legume provides high protein levels, and the grass is persistent and high-yielding. A mixture will also make production more uniform over the growing season. Forage plants do not yield consistently throughout the year and, by mixing those that give high early yields with those that do well later, pasture production is stabilized. Forage may be conserved for winter use as hay (that is, dried) or as silage. In making silage the sugars from the forage are used to produce lactic acid, which pickles the forage.

Fertilizing. When a forage crop is harvested for hay or silage, the whole plant, with the exception of the crown and the root, is removed. This involves a substantial loss of soil nutrients and may be contrasted with grazing when much of the nutrient material in the forage is returned directly to the pasture in the animal's excrement. A grass and legume hay crop will remove about 80 kg (176 lb) of nitrogen, 8 kg (17 lb) of phosphorus, and 33 kg (73 lb) of potassium, if the annual yield level is about 4 metric tons per hectare (8 tons/acre). Fertilizer applications for such a crop are most important if soil fertility is not to be drastically reduced. *See* FERTILIZER.

Applications of nitrogen are essential for the grasses grown alone. Grasses require large amounts of nitrogen, which they take up in the form of nitrate ions. For legumes, which fix their own nitrogen, potassium is an essential element. It enhances legume yield, quality, longevity, and winter survival on many soils.

Plant health. Forage plants are affected by diseases, insects, and, most seriously, weeds. Weeds play a dominant role because they compete for light, water, and nutrients, and also because they harbor insects and disease organisms. Since pastures typically are mixtures of perennial broad-leaved and grassy plants, they present many unique weed control problems. The destruction of weeds by cultivation, which takes place every spring and fall in arable fields, is a rare and very expensive event on grazing lands, and so selective herbicides are commonly used. These substances are plant hormones that kill weeds and leave the forage unharmed when applied at appropriate rates. However, the most desirable weed control is obtained by a close soil cover of forage plants. Then the weeds have no opportunity to establish. *See* HERBICIDE.

The insects that attack forage plants may feed on their foliage, suck out their sap, or, in the form of underground grubs, eat their roots. Control may be chemical or biological. Concern for the environment has led to the use of other insects or diseases to control pests and to breeding resistant cultivars. *See* INSECT CONTROL, BIOLOGICAL.

Bacteria, fungi, viruses, and mycoplasma all cause diseases in legumes and grasses. While chemical control methods exist, forages are not highly priced crops, and the most desirable disease protection is by breeding resistant cultivars. The most serious diseases of forages are those that attack at the seedling stage. At that time, plant populations may be so seriously reduced that the whole field has to be replanted. The application of a fungicide to the seed before planting frequently gives protection during establishment.

Pasture management. Skill and experience in the evaluation of field situations are important to the pasture manager. There are two major tools at the disposal of a manager: the system to be used (number of fields and duration of use) and the stocking rate (number of animals per unit of land). Most grazing systems include a period when livestock are excluded from the pasture and the plants are allowed to regrow. There must be compromises in the choices the manager makes. Pasture production varies during the growing season. In spring, plant growth may exceed animal needs. In midsummer and fall it is slow. The animal needs must always be met without overgrazing or damaging the pasture. Also, optimum animal weight gains per area of land are not achieved by the same stocking rates as are the most favorable gains per individual animal. These factors must be balanced against a background of large variations in forage yield due to climatic changes within and between seasons.

P. D. Walton

Bibliography. B. N. Chatterjee and P. K. Das, *Forage Crop Production*, 1989; A. Dorrat, Y. Cohen, and A. Goldman, *Irrigated Forage Production*, 1993; G. Fahey et al. (eds.), *Forage Quality, Evaluation, and Utilization*, 1994; M. E. Heath, R. E. Barnes, and D. S. Metcalfe, *Forages: The Science of Grassland Agriculture*, 4th ed., 1985; P. D. Walton, *Principles and Practices of Plant Science*, 1988.

Foraminiferida

An order of Granuloreticulosia in the class Rhizopodea. Foraminiferans are dominantly marine protozoa (single-celled animals), with a secreted or agglutinated shell, or test, enclosing the continually changing ameboid body that characterizes this and other orders of the superclass Sarcodina. Their unique combination of long geologic history, ubiquitous geographic distribution, and exceptional diversity of test composition, form, and structure make the foraminiferans the most useful of all marine fossils for stratigraphic correlation, geologic age

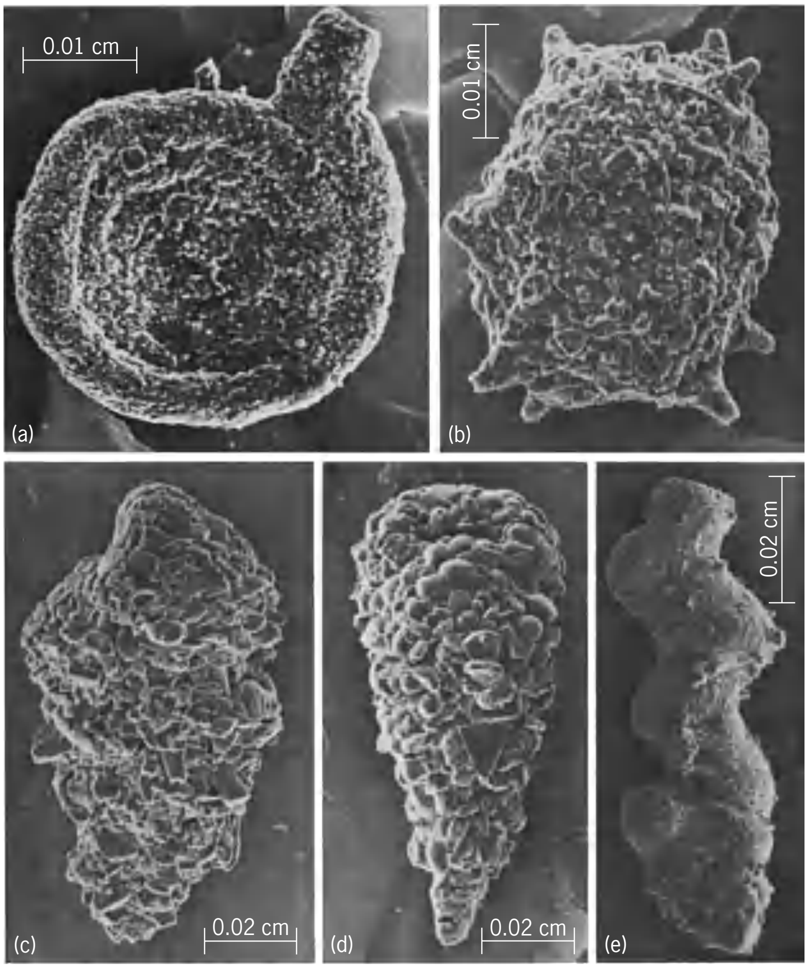


Fig. 1. Scanning electron micrographs of foraminiferans of suborder Textulariina. Superfamily Ammodiscacea: (*a*) *Ammodiscus*, a planispirally enrolled, fine-grained agglutinated test, from Silurian of Missouri. Superfamily Astrorhizacea: (*b*) *Thurammina*, a single globular chamber with apertures on short necks, from Silurian of Missouri. Superfamily Textulariacea: (*c*) *Textularia*, from Holocene of Timor Sea; side view, and (*d*) edge of biserial test coarsely agglutinated of shell fragments. Superfamily Verneuilinacea: (*e*) *Gaudryinella*, with triserial base followed by biserial and zigzag uniserial chambers, wall finely agglutinated, from Upper Cretaceous of Texas. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

dating of sediments, and paleoecologic interpretation. Although small in size, commonly less than 0. 04 in. (1 mm), with a range between 20 micrometers and 4. 8 in. (12 cm), foraminiferans may be abundant, more than 2600 living specimens having been recorded on 1. 5 in.^{2} (10 cm^2) of sea floor. Their tests accumulated in great numbers and are recoverable from small quantities of sediment, rock outcroppings, well cores or cuttings, or ocean dredging and submarine coring.

Living animal. Temporary extrusions (pseudopodia) from the ameboid body form a delicate anastomosing network 2–10 times the test diameter. The pseudopodial net may arise solely from the apertural region of the test in those with tectinous, porcelaneous, and agglutinated walls, or may radiate in all directions through many tiny perforations of the hyaline test wall. The pseudopodia variously serve in capture, ingestion, and digestion of food, in test and temporary cyst construction, for anchorage, and for locomotion. The finely granular cytoplasm consists of darker-colored endoplasm surrounded by a thin, transparent outer layer of ectoplasm. The characteristic granular streaming of the continuously moving cytoplasm differentiates the Foraminiferida from the other Sarcodina orders: the Amoebida, Arcellinida, and Gromiida.

Food. Diatoms are the major food of foraminiferans, although such other organic material as bacteria, phytoflagellates, fragments of larger algae, ciliates, radiolarians, small crustaceans, nematode worms, and other foraminiferans may be used. The food may be carried into the test or may be digested by the external pseudopodia and only the dissolved nutrients ingested. Symbiotic algae have been reported in both benthic and planktonic species. Foraminiferans are in turn used as food by various detritus feeders on the sea floor.

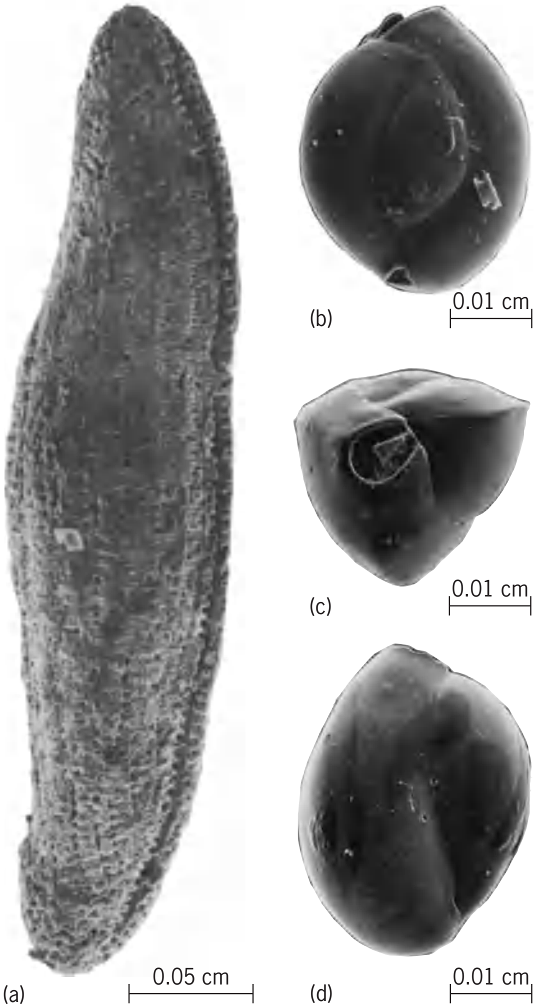


Fig. 2. Scanning electron micrographs of foraminiferans of suborder Miliolina. Superfamily Alveolinacea: (*a*) *Alveolinella*, an isomorph of the microgranular fusulines, but with porcelaneous wall; abraded surface shows the many tiny chamberlets; apertural face at right with numerous apertural pores; from Holocene of Timor Sea. Superfamily Miliolacea: *Triloculina*, from Holocene of Adriatic Sea: (*b, d*) views of opposite sides, each chamber one-half coil in length with coiling in three planes, and (*c*) top view, aperture on final chamber, with bifid tooth projecting from one margin. Wall is porcelaneous and imperforate. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

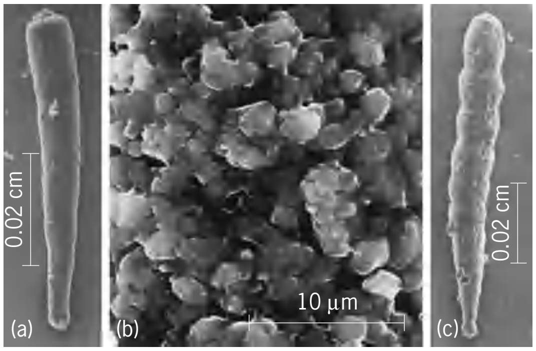


Fig. 3. Scanning electron micrographs of foraminiferans of suborder Fusulinina. Superfamily Eardlandiacea: *Earlandia*, from Pennsylvanian of Texas: (*a*) side view of elongate, undivided tubular test; (*b*) surface of wall, which is composed of tiny angular calcite grains randomly packed. Superfamily Nodosinellacea: (*c*) *Earlandinita*, showing uniserial chambers; from Pennsylvanian of Texas. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

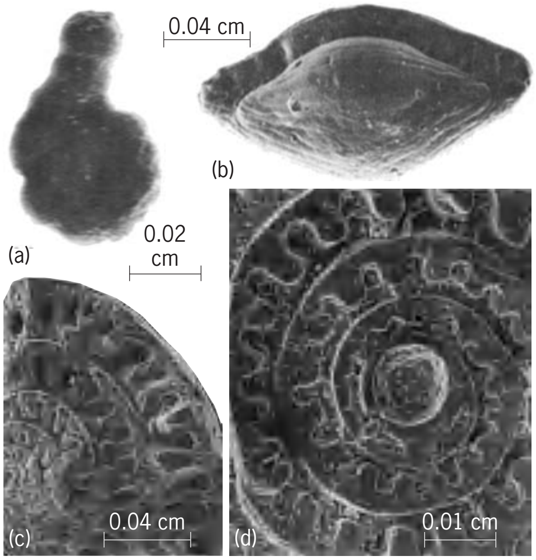


Fig. 4. Scanning electron micrographs of foraminiferans of suborder Fusulinina. Superfamily Endothyracea: (*a*) *Endothyranella*, early chambers enrolled, later ones uncoiled, from Pennsylvanian of Texas. Superfamily Fusulinacea: *Triticites*, from Pennsylvanian of Texas: (*b*) exterior of fusuline test, showing elongate axis of coiling; (*c*) part of transverse weathered half-section showing alveolar wall, proloculus at left; and (*d*) early portion of transverse section enlarged, showing globular proloculus in center, thick wall, and many small chambers per whorl. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

Reproduction and dimorphism. Two methods of reproduction normally alternate, an asexually produced haploid generation (gamont) giving rise by sexual reproduction to a diploid generation (agamont or schizont), which in turn by asexual multiple fission immediately after meiosis produces the new haploid generation. Distinct but morphologically similar haploid and diploid generations occur in the animal kingdom only in the foraminiferans. The entire parent individual is involved in both asexual schizogony and in gamogony; nothing but an empty test remains of the old generation after reproduction.

The separate generations also can be differentiated by test characters. In fact, dimorphism was observed in fossils long before its ontogenetic basis was known. The uninucleate gamont generation normally has a relatively large (megalospheric) first test chamber, or proloculus, whereas the multinucleate agamont generation has a smaller (microspheric) proloculus but a larger, more differentiated adult test. All nuclei may be alike in form and function (homokaryotic), or vegetative and generative nuclei may be differentiated (heterokaryotic), with only the generative nuclei taking part in reproduction. Megalospheric and microspheric tests may be otherwise nearly identical in appearance or sufficiently different that some have been described as separate species. Dimorphism is most pronounced in the more highly evolved calcareous species; the distinctive float chamber formed prior to gametogenesis by *Rosalina* gamonts led to description of a separate genus, *Tretomphalus*.

Modifications of the usual asexual reproduction include shell fragmentation and simple binary fission in some agglutinated types (*Bathysiphon*) or budding and fission (*Halyphysema*). Either asexual schizogony or sexual gamogony may be repeated successively before intervention of the other process to complete the cycle. In sexual reproduction, countless tiny biflagellate gametes typically are released into the water (*Boderia*, *Iridia*, *Peneroplis*, *Elphidium*). Gametes of the Miliolina have a distinctive axostyle. *Glabratella* (superfamily Glabratellacea) has triflagellate gametes, produced by two parent individuals that attach by their umbilical surfaces (plastogamy) to form a joint "brood chamber" in which fertilization occurs. A few large ameboid gametes are produced by two or more parent individuals

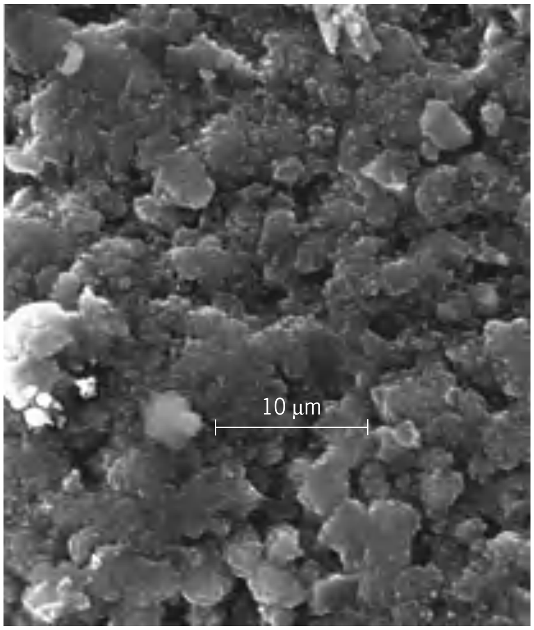


Fig. 5. Enlarged portion of Fig. 4*b* showing surface of microgranular wall constructed of fine, equidimensional calcite granules. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

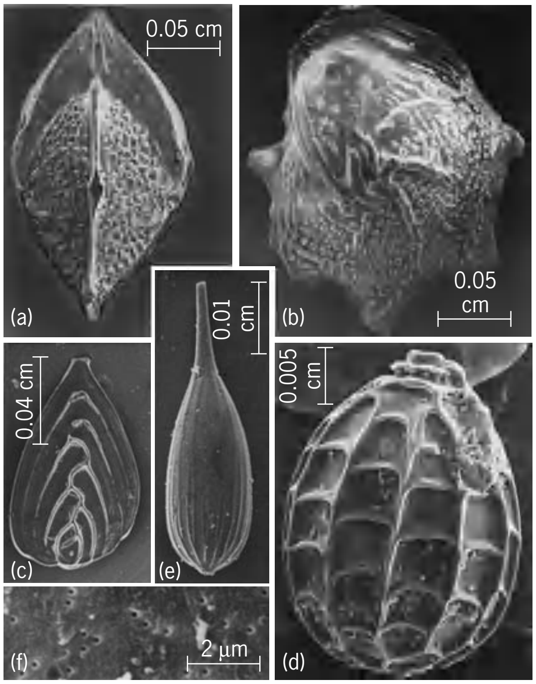


Fig. 6. Scanning electron micrographs of foraminiferans of suborder Lagenina, superfamily Nodosariacea. *Lenticulina*, from Pliocene of Italy: (*a*) edge view showing radiate aperture at top of figure; (*b*) side view of planispirally enrolled test, with surface ribs, nodes, and peripheral keel. (*c*) *Kyphopyxa*, flattened test, with early biserial chambers and later chevron-shaped uniserial ones; aperture at top of figure; from Upper Cretaceous of Texas. (*d*) *Favulina*, single-chambered test with surface costae, from Pleistocene of California. *Laqena*, from Pleistocene of California: (*e*) single-chambered test with elongate apertural neck; (*f*) surface view, showing very small pores in the hyaline wall. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

that in association form a joint reproductive cyst (*Spirillina*, *Patellina*). The large gametes and hence large zygote result in a larger proloculus in the schizont, in contrast to the usual smaller microspheric schizont.

Many other variations in the basic life cycle are known in the approximately two dozen species thus far studied, representing only a few families or genera; other modifications also are probable, for nothing is yet known as to sexual reproduction for many superfamilies, and even some entire suborders (Lagenina, Silicoloculinina, or Robertinina: *Rubratella* may belong to the Robertinina, but no information is yet available as to its wall characters.)

Test. The foraminiferan constructs its own test. Growth of the individual may cease after test construction, or the test may enlarge by continued growth in one or more directions (tubular or branching tests) or by periodic formation of separate but always interconnected chambers. Growth may be either continuous or periodic and may be in a straight line, a planispiral or trochospiral coil, a cycle, or a zigzag. The latter zigzag is expressed by chambered forms in biseriality. Variations and combinations of these growth patterns occur repeatedly in different lineages (isomorphism).

Continual circulation of protoplasm within and without the test is through test openings. A major opening on the final chamber, the aperture, is characteristic in form and position. Supplemental apertures, complex canal and stolon systems, and tiny wall perforations also may occur. The test wall may be thin and delicate or thick and firm, may be simple and homogeneous or have two or more similar or differing layers, and may have a lamellar or labyrinthic structure. The wall may be organic, agglutinated, calcareous, or rarely siliceous.

Tectinous and agglutinated walls. The suborder Allogromiina has an organic (tectinous) test of protein and an acid mucopolysaccharide. Agglutinated walls, characteristic of the suborder Textulariina, consist of foreign particles, quartz and other mineral sand or silt grains, sponge spicules, and tiny shells or shell fragments held in an organic or calcareous cement (**Fig. 1**). Some show selectivity in the foreign material used for test construction and in its arrangement, whereas others utilize a variety of materials in a heterogeneous arrangement. The organic cement is a protein-carbohydrate, at least in part an acid mucopolysaccharide (“tectin”), and morphologically may consist of closely packed strands covering the surface of the cemented grains and filling the intergrain spaces. In others the grains are cemented by a fibrous meshwork, or the organic material may form a foamlike mass of organic bubbles. An inner

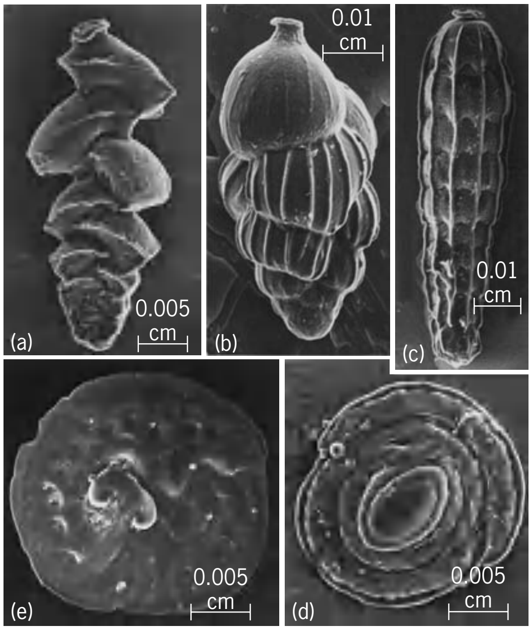


Fig. 7. Scanning electron micrographs of foraminiferans of suborders Rotaliina and Spirillinina. Superfamily Eouvigerinacea: (*a*) *Eouvigerina*, elongate test, early stage biserial, later uniserial, aperture terminal with everted lip; from Upper Cretaceous of Texas. Superfamily Buliminacea: (*b*) *Uvigerina*, elongate triserial test; from Miocene of California. (*c*) *Siphogenerinoides*, reduced early biserial stage, later uniserial; from Upper Cretaceous of Texas. Suborder Spirillinina: *Patellina*, from Pleistocene of California; (*d*) umbilical view of low conical test, showing aperture; (*e*) spiral view showing oval proloculus, enrolled tubular chamber, and later biserial chambers with secondary partitions of later chambers resulting in a crenulate margin. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

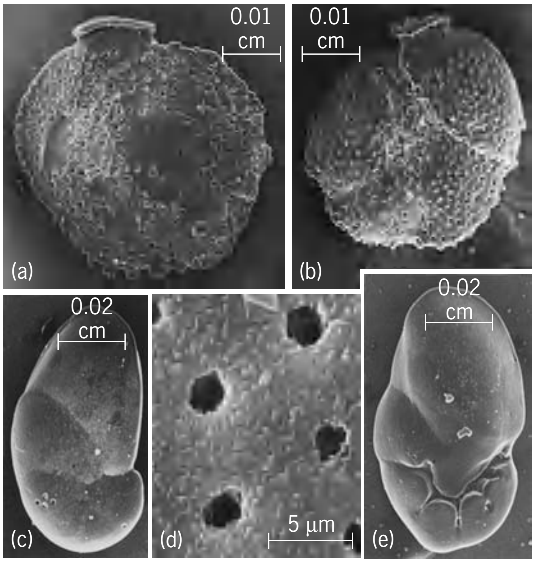


Fig. 8. Scanning electron micrographs of foraminiferans of suborder Rotaliina, superfamily Siphoninacea. *Siphonina*, from Upper Eocene of Mississippi: (*a*) spiral and (*b*) umbilical views of coarsely perforate, trochospiral test with terminal aperture and everted lip. Superfamily Discorbacea: *Cancris*, from Pliocene of Italy: (*c*) spiral and (*e*) umbilical views of finely perforate right- and left-coiled specimens, respectively, the latter showing large imperforate flap that extends over the basal aperture; (*d*) an enlarged part of the spiral side, showing pores and the surface of small subrhombic calcite crystals. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

organic lining separates the test from the cytoplasm, and some also have an outer organic layer.

Although ferruginous, siliceous, and calcareous cements have been reported, the ferruginous material was shown to have consisted of coatings on the foreign particles before their incorporation in the wall, and siliceous cement has not yet been proven to occur. Calcareous cement does occur, particularly in members of the Textulariacea, where it may appear as tiny (0.1-μm) equidimensional grains of low magnesium calcite that may be aligned in rods or bundles of rods that in turn are randomly arranged in the wall. The calcite crystals are enveloped by organic material, suggesting intracellular calcification and later transport with invesicles to the final position of the crystals in the wall. Many of those with calcitic cement (Eggerellidae, Textulariidae, Pseudogaudryinidae, Valvulamminidae, Valvulinidae, Chrysalidinidae, Glaucoamminidae) also have wall pores or canaliculi, which may be closed at each end by an organic plug or lining or may be obscured at the surface by agglutinated material but do not permit passage of the cytoplasm.

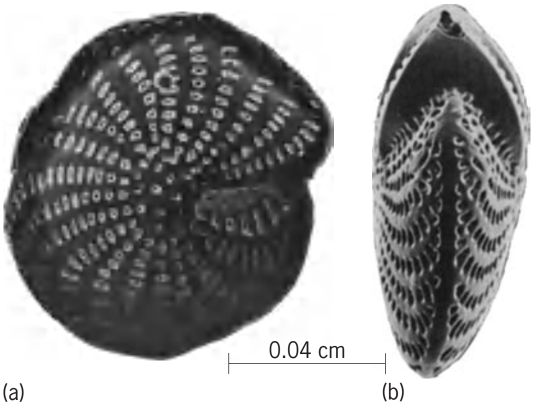


Fig. 10. Scanning electron micrographs of *Elphidium*, suborder Rotaliina, superfamily Rotaliacea, from Holocene of Adriatic Sea. (*a*) Side view showing large sutural pores that open into canal system. (*b*) Edge view with series of apertural pores at lower margin of final chamber face and imperforate peripheral band. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

Porcelaneous walls. The imperforate porcelaneous test (suborder Miliolina) is constructed of randomly arranged needles of high-magnesium calcite, containing considerable organic matter (**Fig. 2**). The three-dimensionally random arrangement of the crystals refracts light in all directions, giving the test a milky opacity. In transmitted light it appears brown. A

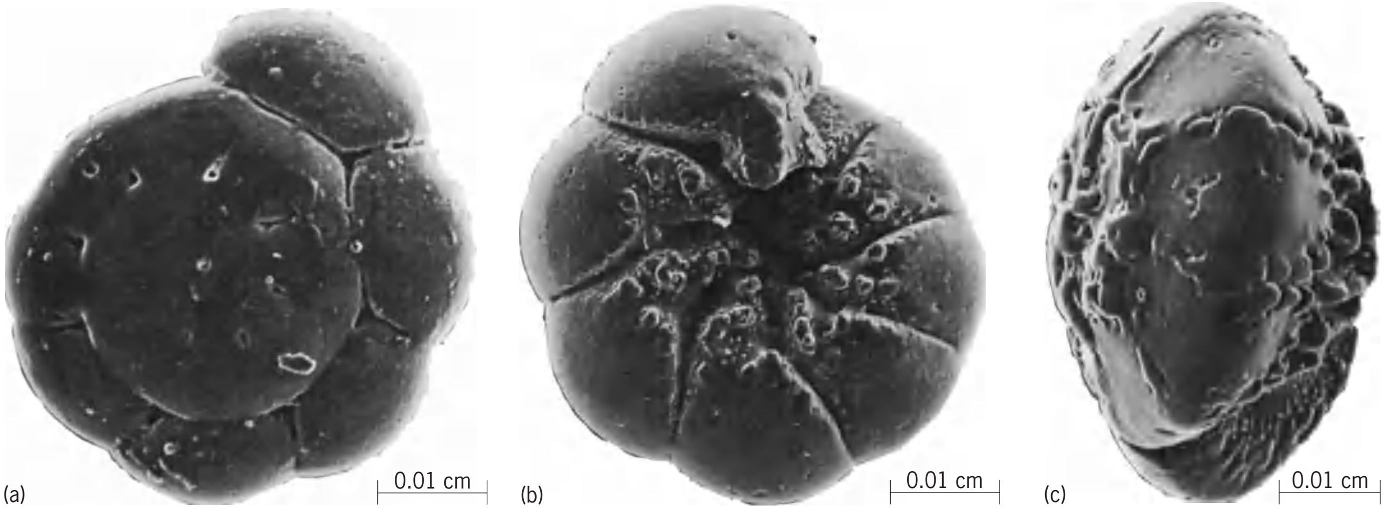


Fig. 9. Scanning electron micrographs of *Ammonia*, suborder Rotaliina, superfamily Rotaliacea, from Pliocene of Italy. (*a*) Spiral, (*b*) umbilical, and (*c*) edge views of trochospiral test showing the umbilical and the sutural fissures and canals, the umbilical pillars, and the basal aperture of the last chamber. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

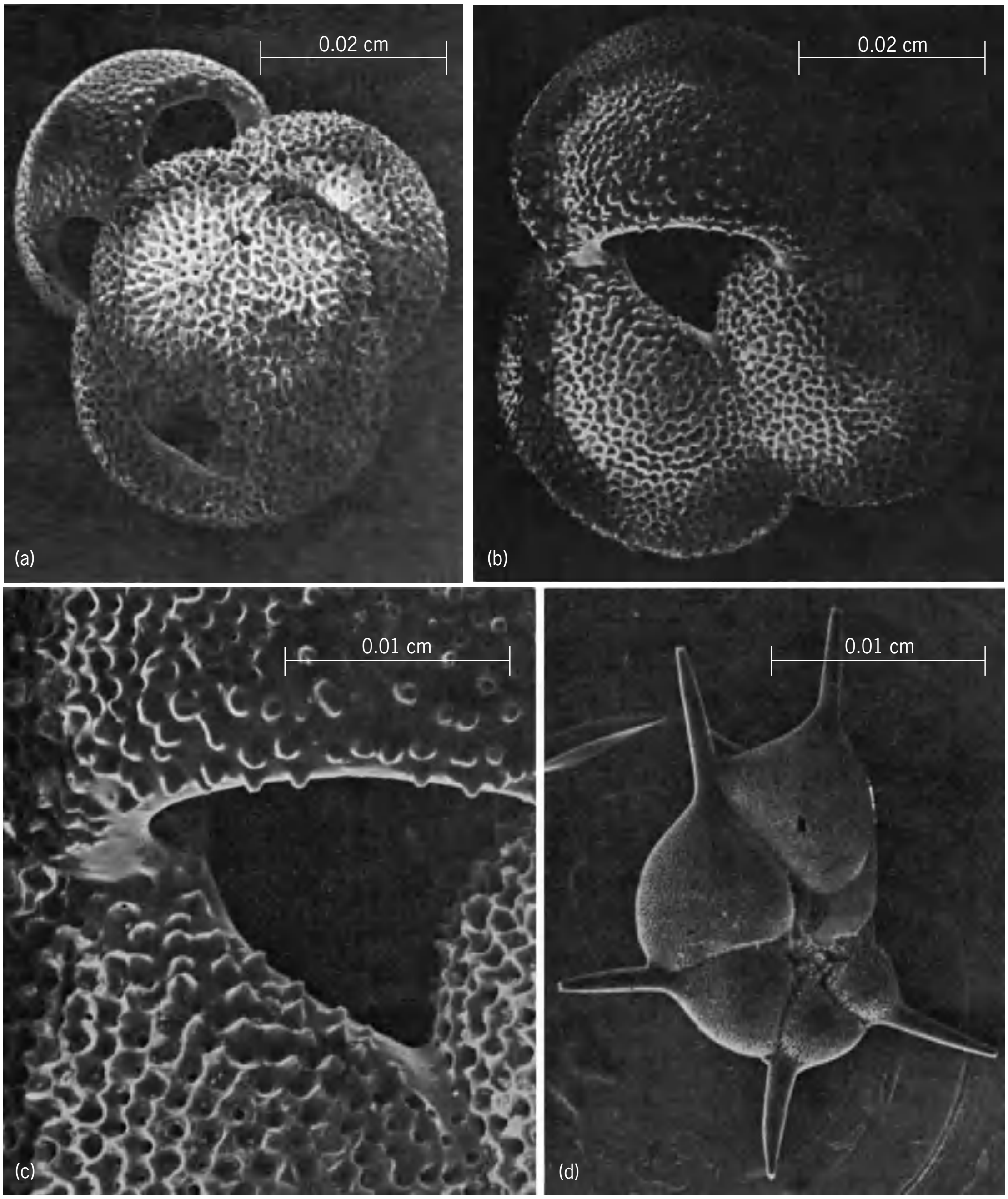


Fig. 11. Scanning electron micrographs of foraminiferans of suborder Globigerinina, superfamily Globigerinacea. *Globigerinoides*, from Holocene of Caribbean Sea: (*a*) spiral and (*b*) umbilical views of the thin-walled, coarsely perforate test of the planktonic species; (*c*) enlargement of apertural region of *b* showing imperforate apertural margin and coarse perforations elsewhere; short, blunt spines are the broken ends of very fine elongate delicate spines that project radially during life. Superfamily Hantkeninacea, from Upper Eocene of Mississippi. (*d*) *Hantkenina*, showing perforate chamber wall and imperforate tubulospines and apertural flap of last chamber at center right. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

surface coating may have parallel groups of calcite laths. True pores are lacking, but pseudopores may occur in the surface layer; these end in the randomly oriented layer below and do not completely penetrate the wall.

Microgranular calcite. The Fusulinina have a microgranular wall of tiny, tightly packed equidimensional subangular calcite crystals (**Figs. 3–5**). The granules may be regularly aligned, producing a fibrous appearance; as many as four separate granular and fibrous wall layers may occur in various combinations. Some Fusulinacea have perforated and even alveolar walls. Other characteristic structures result from solution and secondary deposits formed in the enlarging test.

Lamellar hyaline tests. Hyaline calcareous walls occur in the suborders Lagenina, Spirillinina, Robertinina, Globigerinina, and Rotaliina (**Figs. 6–15**). Successive laminae are added with growth. A thick organic lining, probably a polysaccharide, is just within the calcareous wall; constructed first, it gives form

Fig. 12. Scanning electron micrograph of *Spiroplecta*, suborder Globigerinina, superfamily Heterohelicacea, showing biserial test and inflated chambers and weakly striate surface; from Upper Cretaceous, Texas. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

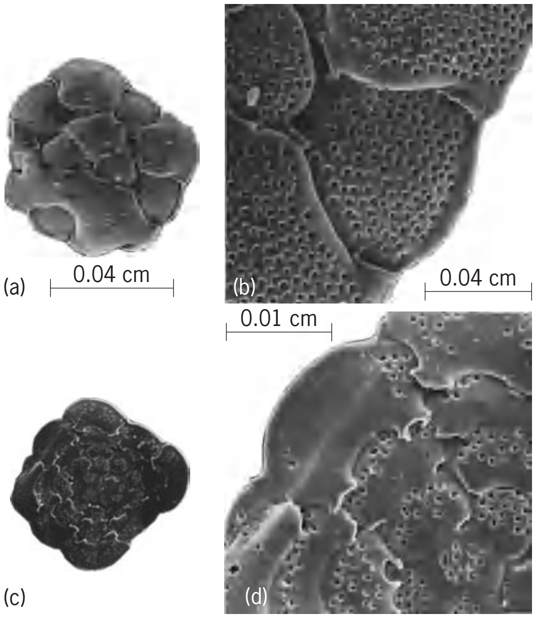


Fig. 13. Scanning electron micrographs of *Planorbulina*, suborder Rotaliina, superfamily Planorbulinacea, test with trochospiral early stage and proliferating later stage; from Pliocene of Italy. (*a*) Umbilical view; (*b*) enlargement, showing large pores on surface and two or more large apertures at the edge of each chamber of the outer whorl; (*c*) view of attached spiral side; (*d*) enlargement of *c* showing fewer perforations and small apertural projections from each later chamber. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

and shape to the newly added chamber. Electron microscopy has detected a proteinaceous basal membrane just outside the lining that serves as the organic matrix for calcification. As each new chamber is added, its new basal membrane also extends over the previously formed test; a new calcareous lamina is added to all exposed surfaces of previous chambers as the newly formed one is calcified. The test lamellae are not merely juxtaposed calcite layers, but the calcite layers are interspersed with fine separating membranes. Foraminiferans with numerous chambers in each whorl develop very thick walls, layer by layer with each chamber addition (as in *Nummulites*), whereas those with few chambers per whorl remain thin-walled (*Globigerina*). The inner organic lining is also added against the previous apertural face and, within earlier chambers, continually thickening with new chamber additions. The resultant thick inner lining of early chambers may give them a distinct brown coloration in the living test, as in *Rosalina*.

Lamellar septa. Depending upon the method of its formation, the new test extension or new chamber wall may consist of a single layer (Textulariina, Involutinina, Spirillinina, Miliolina, Lagenina) or be double, because of calcification at both sides of the organic membrane serving as a template for the new chamber. As the next chamber is added, the previous apertural face becomes an intercameral septum, retaining the original monolamellar or bilamellar character, whereas the outer test wall may be enveloped by an additional lamella at each instar. The originally bilamellar septa of the Rotaliacea and Nummulitacea are additionally thickened by a septal flap attached to the previous apertural face as the new chamber is formed. Spaces left between this septal flap and the septum may form a complex system of radial canals, umbilical cavities, and intraseptal and subsutural canals in these groups.

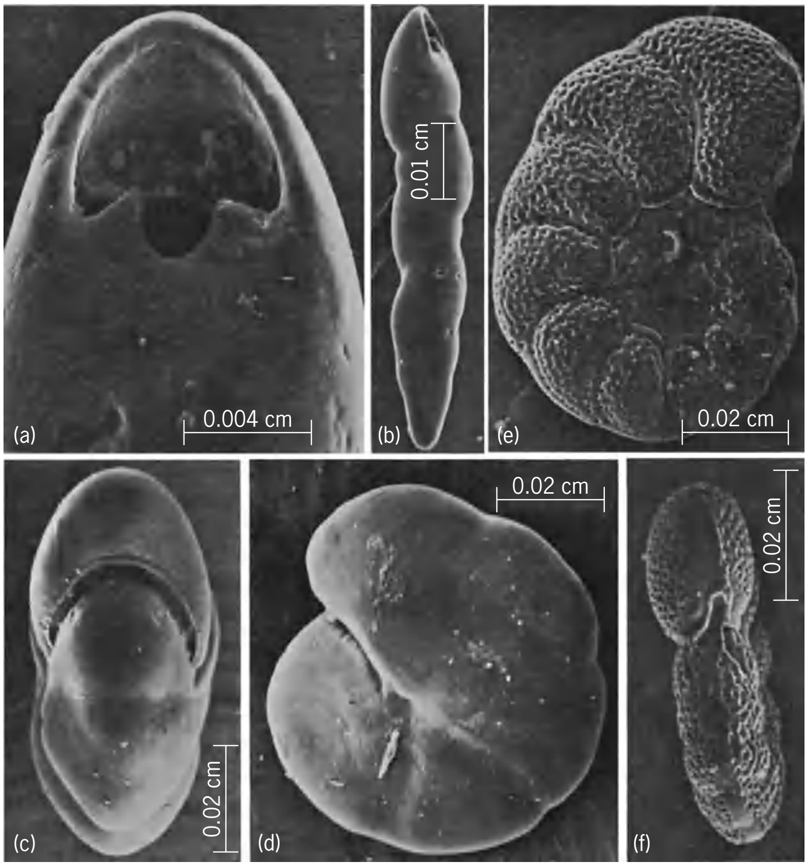


Fig. 14. Scanning electron micrographs of foraminiferans of suborder Rotaliina. Superfamily Pleurostomellacea: *Pleurostomella*, from Lower Cretaceous of England: (*a*) apertural view of final chamber showing two projecting teeth; (*b*) side view of biserial to uniserial test with perforate granular wall and hooded aperture. Superfamily Nonionacea: *Pullenia*, from Holocene of Antarctica: (*c*) apertural and (*d*) side views of planispiral finely perforate test with granular wall. Superfamily Chilostomellacea: *Holmanella*, from Miocene of California: (*e*) spiral and (*f*) edge views of bievolute, planispiral test, with a coarsely perforate optically granular wall, imperforate peripheral margin, and slitlike aperture extending up the terminal face. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

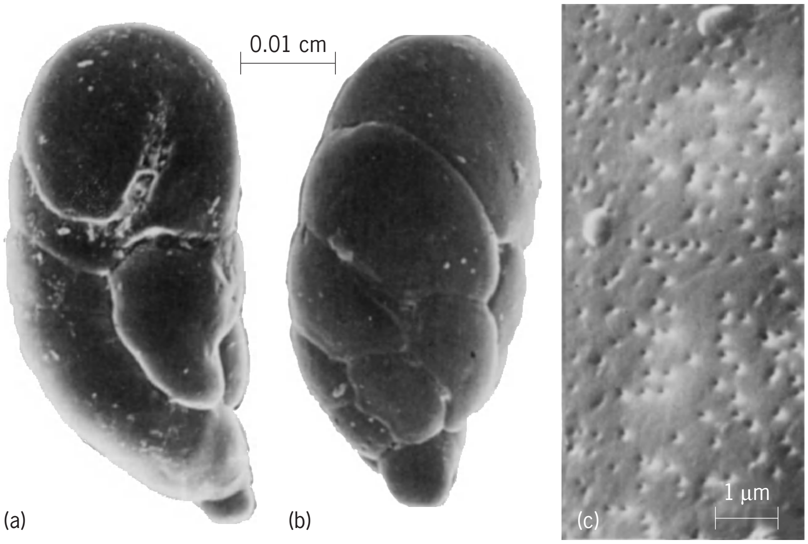


Fig. 15. Scanning electron micrographs of *Robertina*, suborder Robertinina, superfamily Robertinacea, from Pleistocene of California. (*a*) Apertural and (*b*) opposite sides of test with median subdivisions of the chambers; (c) surface with very fine pores in optically radial aragonitic wall. (*Courtesy of R. B. MacAdam*, *Chevron Oil Field Research Co.*)

Pores. Wall perforations of the hyaline foraminiferans may be extremely tiny or relatively large (Figs. 6*f*, 8, 10, 11, 13, 14, and 15). Two sizes may be interspersed; small pores may enlarge, coalesce, or become closed with lamellar growth; and some hyaline species seem completely lacking in perforations. Pores may be randomly spaced or restricted to certain areas of the chambers. Apertural lips, tooth plates, and surface ribs and carinal bands are commonly imperforate. Pores may contain organic plugs at the level of successive wall lamellae; perforated plugs are regarded as having a filtering function, but others are solid. Pores also may penetrate the calcite wall but not the organic lining and hence be nonfunctional as passages. Pores are also organically lined and may be only a remnant of the organic matrix, strengthening the test by binding the organic membranes of successive lamellae.

Radial and granular structure. Two main types of hyaline microstructure occur, termed optically radial (or hyaline radial) and optically granular (hyaline oblique), in reference to the orientation of the calcite or aragonite crystals in the wall. In the optically radial structure, each growth lamella consists of many layers of subrhombic calcite crystals, whose preferred orientation, with *c* axes normal to the test surface, gives a characteristic extinction pattern (black cross with concentric rings of color) under crossed Nicol prisms in the polarizing microscope. This occurs in the Spirillinina, Lagenina, Globigerinina, and many Rotaliina. Walls of the Robertinina are similarly constructed of the orthorhombic aragonite form of calcium carbonate. Crystals in the wall of Lagenina have irregular outlines, the surface ultrastructure showing irregularly formed interlocking units, with boundaries marked by a honeycomblike organic matrix. Crystal boundaries are irregular, but their preferred orientation is like that of other hyaline radial walls. In Spirillinina the calcite of the entire test is in optical continuity, acting as a single crystal in polarized light.

The optically granular or hyaline oblique wall (Fursenkoinacea, Delosinacea, Pleurostomellacea, Annulopatellinacea, Chilostomellacea, most Nonionacea, some Loxostomatacea, Cassidulinacea, Turrilinacea, and Siphoninacea, and a few Rotaliacea) contains smaller calcite grains, with preferred crystal orientation on a rhombohedral face, so that the *c* axis is oblique to the surface instead of perpendicular. In the polarizing microscope the granular crystals show only tiny flecks of color instead of a black cross.

In addition, two other types of wall structure have been described, an intermediate structure and a compound structure. The intermediate structure occurs in some Planorbulinacea (*Cibicides*, *Planorbulina*) and Chilostomellacea (*Gyroidina*). In this structure the crystal units are small and do not extend continuously through the wall. Near the outer surface the *c* axes are nearly normal to the wall surface as in the radial structure, but in the interior the crystal units are oblique to the surface as in the granular type. Hence, under polarized light the extinction pattern shows characteristics of both radial and granular walls.

The compound structure found in *Lobatula* (superfamily Planorbulinacea) consists of a wall with separate layers of differing type. An inner layer has *c* axes perpendicular to the test surface as in the radial type, but the outer primary layer has *c* axes and one edge of each microcrystal parallel to the test surface, the combined layers resulting in a granular extinction pattern under polarized light.

Spicular structure. The wall of the suborder Carterinina contains many elongate rodlike or fusiform spicular crystals whose *c* axes lie lengthwise of the spicules. The spicules or aligned bundles of spicules are held in a groundmass of similar smaller spicules, an inner layer being aligned normal to the test periphery and a surface layer being aligned parallel to the test surface and periphery. Later chambers have progressively larger spicules, whose lamellar structure indicates their further enlargement within the test.

Classification. As is true of most protists with skeletons or tests, systematic differentiation and classification of foraminiferans is based on test composition, microstructure, and gross morphology. Information available concerning cytoplasmic characters, life cycles, and so on has shown good agreement with this classification, although the function and origin of many shell characters believed to be of systematic importance (canal systems, pores, septal doubling, and apertural tooth plates) are not all understood. Characters of suborders and superfamilies are given in the **table**.

Evolutionary development. Foraminiferans have shown a remarkable diversification and rapid evolutionary development over their 550-million-year known history. Some with uncomplicated tests have been reported to have a long geologic range, but most with diagnostic shell structure and morphology had relatively short histories, each successively

Characteristics of foraminiferan suborders and superfamilies

Superfamily	Test wall	Test form	Distribution
Suborder Allogromiina	Organic		
Lagynacea (5 families, 56 genera)	Membranous to tectinous; surface may be agglutinated or ferruginous	Single ovid, tubular, or irregular chamber, rarely multilocular; free or attached; one or more apertures	Marine, brackish, fresh water; benthic, a few planktonic; Cambrian to Holocene
Suborder Textulariina	Agglutinated with organic or calcareous cement		
Astrorhizacea (12 families, 93 genera)	Simple to labyrinthic; with organic lining and organic cement	Irregular, rounded, tubular or branching nonseptate chamber, or cluster of chambers; one or more apertures	Marine, benthic; Cambrian to Holocene
Komokiacea (2 families, 7 genera)	Silt particles in organic cement; organic lining	A system of branching tubules; aperture consists of wall pores	Marine, benthic; abyssal and hadal; Holocene
Hippocrepinacea (3 families, 18 genera)	Of quartz particles and sponge spicules; organic cement	Proloculus and undivided tubular or flaring second chamber; free or attached; aperture terminal, single	Marine, benthic; Ordovician to Holocene
Ammodiscacea (1 family, 31 genera)	Arenaceous; organic cement	Proloculus and enrolled to uncoiled, tubular, nonseptate second chamber; aperture terminal, single	Marine, benthic; Cambrian to Holocene
Rzehakinacea (1 family, 11 genera)	Finely agglutinated; organic cement	Enrolled, two to three chambers per whorl; planispiral or coiled in varied planes	Marine, or brackish to saline lakes; benthic; Cretaceous to Holocene
Hormosinacea (6 families, 32 genera)	Firmly cemented and agglutinated; thin with single layer of grains to very thick; may be perforated by coarse tubuli	Multilocular, chambers uniserial to branching; free or attached; one or more apertures	Marine, bathyal to abyssal; Ordovician to Holocene
Lituolacea (9 families, 62 genera)	Agglutinated on organic base; no pores	Multilocular; free or attached; planispiral to uncoiled; aperture single to multiple, basal to terminal	Brackish to marine, benthic; neritic to abyssal; Devonian to Holocene
Haplophragmiacea (7 families, 25 genera)	Agglutinated, solid to alveolar, rarely with canaliculate septa	Streptospirally enrolled to uncoiled, may branch; chambers simple or with inner partition; aperture basal to terminal	Brackish to marine, benthic; neritic to bathyal; Jurassic to Holocene
Biokovinacea (3 families, 11 genera)	Finely agglutinated, imperforate outer layer; inner part may be perforate or canaliculate	Planispiral or may uncoil; aperture single to multiple, basal to areal	Benthic, marine, on calcareous substrate; Jurassic to Upper Cretaceous
Coscinophragmatacea (2 families, 5 genera)	Coarsely perforate to alveolar; pores closed by organic inner lining of wall	Attached, coiled to uncoiled or branching; terminal aperture, rounded to irregular	Benthic, marine; Triassic to Holocene
Cyclolinacea (2 families, 10 genera)	Microgranular, calcareous, agglutinated; with imperforate epidermis	Discoidal; early stage coiled, later peneropliform to cyclic chambers; may have secondary partitions or pillars; aperture multiple, may result in radial or oblique stolon system	Benthic, marine; in calcareous sediments; Jurassic to Upper Cretaceous
Loftusiacea (6 families, 47 genera)	Agglutinated, with imperforate outer layer and alveolar or reticulate subepidermal layer	Multilocular, coiling planispiral to streptospiral, rarely trochospiral; later may uncoil or become biserial; interior simple, or with pillars, beams, and rafters; aperture single to multiple, basal to areal, or terminal	Marine, benthic; Triassic to Holocene
Spiroplectamminacea (5 families, 27 genera)	Solid and noncanaliculate, agglutinated with organic cement	Early planispiral to streptospiral stage, rarely biserial; later may be biserial or uniserial; chambers simple, undivided; aperture single to multiple, basal to areal or terminal	Marine, benthic; upper Carboniferous to Holocene
Pavonitacea (2 families, 7 genera)	Noncanaliculate	Early stage coiled, triserial or biserial, later biserial or uniserial; interior may have vertical beams, rarely rafters; aperture basal to terminal, single to multiple	Marine, benthic; Upper Cretaceous to Pliocene
Trochamminacea (2 families, 32 genera)	Agglutinated, organic cement, noncanaliculate	Low trochospiral coil, or uncoiling; secondary partitions due to infolding of wall; aperture basal to areal, single to multiple, may have supplementary umbilical openings	Benthic, brackish to marine; upper Carboniferous to Holocene
Verneuilinacea (4 families, 39 genera)	Agglutinated, solid, noncanaliculate	Elongate; early stage trochospiral, triserial, or biserial, later may be uniserial and rectilinear; aperture basal, areal, or terminal on neck	Benthic, marine; upper Carboniferous to Holocene
Ataxophragmiacea (8 families, 61 genera)	Noncanaliculate, agglutinated; may be alveolar or have imperforate epidermis and reticulate subepidermal layer	Trochospiral, later reduced to biserial or uniserial, or with increased number of chambers per whorl; internally may have secondary partitions, buttresses, or pillars; aperture basal to areal, simple to cribrate	Benthic, marine; Triassic to Holocene
Orbitolinacea (1 family, 41 genera)	Noncanaliculate, agglutinated	Low to high conical; early stage trochospiral or with protoconch and deuteroconch; later chambers uniserial, discoidal, numerous, with radial and transverse partitions and vertical pillars	Benthic, marine; Jurassic to Oligocene
Textulariacea (7 families, 65 genera)	Agglutinated, calcareous cement, canaliculate	Trochospiral, triserial or biserial in early stage; later biserial or uniserial; chambers simple or with internal partitions or pillars; aperture basal to areal, single to multiple	Benthic, marine; Jurassic to Holocene

(*cont.*)

Characteristics of foraminiferan suborders and superfamilies (cont.)

Superfamily	Test wall	Test form	Distribution
Suborder Fusulinina	Nonlamellar, microgranular calcite, of equidimensional crystals a few micrometers in diameter		
Parathuramminacea (10 families, 44 genera)	Homogeneous, one to three layers	Single globular to irregular chamber or cluster of chambers; free or attached; aperture not recognizable, or multiple, on necks or conical projections	Benthic, marine; Silurian to Permian
Earlandiacea (3 families, 11 genera)	Simple, rarely with some agglutinated material	Globular proloculus and undivided straight or enrolled tubular second chamber; aperture terminal	Benthic, marine; Silurian to Permian
Archaediscacea (2 families, 21 genera)	One or more layers, one with radial structure, one compact; one or both sides may have axial thickenings	Proloculus followed by planispiral, streptospiral, or trochospiral tubular second chamber; aperture at open end of tubular chamber	Benthic, marine; Carboniferous
Moravamminacea (3 families, 12 genera)	Single-layered	Proloculus followed by enrolled or rectilinear tubular second chamber; incipient or partial septa present; aperture terminal	Benthic, marine; Silurian to lower Carboniferous
Nodosinellacea (2 families, 7 genera)	Simple, or with outer granular layer and fibrous inner layer	One or more distinct uniserial chambers, rectilinear to arcuate; aperture terminal	Benthic, marine; Silurian to Permian
Geinitzinacea (2 families, 8 genera)	Two layers, outer one radially fibrous and inner one dark and microgranular; secondarily nonlamellar or partially lamellar	Uniserial; rounded to flattened in section; aperture terminal	Benthic, marine; Devonian to Permian
Colaniellacea (1 family, 4 genera)	Two layers, outer one vitreous, inner one finely granular	Uniserial; chambers strongly overlapping; with internal radial, vertical interseptal partitions; aperture terminal	Benthic, marine; Devonian to Permian
Ptychocladiacea (1 family, 2 genera)	Microgranular, banded, possibly due to transverse tubuli	Attached; uniserial, branching or spreading; no distinct aperture	Benthic, marine; Devonian to upper Carboniferous
Palaeotextulariacea (3 families, 22 genera)	One or two layers; may have agglutinated coating	Biserial, enrolled biserial, or uniserial; advanced ones with secondary vertical partitions; aperture basal to terminal, single to multiple	Benthic, marine; Devonian to Permian
Tournayellacea (2 families, 48 genera)	One or two layers; may have agglutinated inclusions; may have secondary deposits	Proloculus and tubular planispirally to streptospirally coiled second chamber, or may uncoil; incipient septa may produce teardrop-shaped chambers; single to cribrate aperture	Benthic, marine; Devonian to upper Carboniferous
Endothyracea (2 families, 59 genera)	Two to three layers; may have agglutinated particles also; supplementary deposits may occur	Planispiral to streptospiral coiling; later may be uncoiled and rectilinear; true septa present; basal to areal aperture, single or rarely cribrate; may have supplementary sutural openings	Benthic, marine; Devonian to Permian
Tetrataxacea (4 families, 7 genera)	One or two layers	Conical; trochospiral, evolute spiral side and involute umbilical side; few to many chambers per whorl; numerous secondary partitions; aperture umbilical	Benthic, marine; lower Carboniferous to Permian
Fusulinacea (8 families, 167 genera)	One to four layers; resorption may produce foramina or tunnels; secondary deposits form axial fillings, chomata, parachomata, and tectoria	Globular, lenticular, or fusiform; tightly coiled or may later uncoil; many chambers per whorl; septa plane to fluted; advanced ones with secondary septula; aperture a basal slit to multiple openings	Benthic, marine; Devonian to Permian
Suborder Involutinina	Homogeneous, calcareous originally aragonite but commmonly recrystallized		
(4 families, 30 genera)	Perforate, radiate; may be recrystallized to homogeneous microgranular structure; lamellar thickening or pillars on one or both sides	Proloculus followed by planispiral to trochospiral enrolled, undivided tubular second chamber; aperture at open end of tubular chamber	Benthic, marine; in calcareous deposits; Permian to Upper Cretaceous
Suborder Spirillinina	Optically single calcite crystal; formed by marginal accretion		
(2 families, 11 genera)	Of calcite crystals, which may show preferred orientation; pseudopores of umbilical surface filled with organic matter; marginal calcification, not on organic template	Planispiral to high trochospiral coil; proloculus and coiled undivided second chamber or with two to three chambers per whorl; may have secondary radial partitions; aperture at open end of tube on periphery or directed toward umbilicus	Benthic, shallow marine; Triassic to Holocene
Suborder Carterinina	Rodlike spicules secreted by the organism		
(1 family, 1 genus)	Inner organic lining, rodlike or fusiform spicules, each a single low-magnesium calcite crystal oriented with respect to periphery, held in organic matrix	Attached; trochospiral early chambers becoming crescentic and finally irregular, later ones with secondary radial beams; single umbilical aperture, later becoming multiple	Benthic, marine, shallow tropical water; Eocene, Holocene

Characteristics of foraminiferan suborders and superfamilies (cont.)

Superfamily	Test wall	Test form	Distribution
Suborder Miliolina	Imperforate porcelaneous, of fine randomly oriented rodlike crystals of high-magnesium calcite, containing organic material, may have surface veneer of tabular crystals		
Squamulinacea (1 family, 2 genera)	As in suborder	Single globular to subcylindrical chamber; free or attached	Benthic, marine; Upper Cretaceous, Holocene
Cornuspiracea (7 families, 73 genera)	As in suborder	Free or attached; planispiral, streptospiral, trochospiral, uncoiled, or irregular, spreading to discoidal; proloculus and undivided second chamber, later chambers may be tubular, rounded, or cyclic, and may be subdivided; aperture terminal or row of pores	Benthic, marine; Lower Carboniferous to Holocene
Miliolacea (6 families, 113 genera)	As in suborder; wall may have pseudopores	Enrolled, two or rarely three tubular chambers per whorl, added in various planes around elongate axis; may have longitudinal or transverse secondary subdivisions; aperture terminal, single or multiple	Benthic, marine, mostly warm water; Triassic to Holocene
Alveolinacea (3 families, 32 genera)	As in suborder; may have basal wall thickening, subepidermal partitions, or pillars	Enrolled about elongate axis; proloculus and flexostyle followed by planispiral, streptospiral or milioline stage; adult may uncoil; may have internal septulae or pillars; aperture multiple, rarely single	Benthic, marine, shallow tropical water; Lower Cretaceous to Holocene
Soritacea (6 families, 59 genera)	As in suborder; early stage coarsely perforate or pitted, later imperforate; some extinct taxa perforate throughout	Planispiral to uncoiling, flabelliform or discoid; may have internal interseptal partitions or pillars; aperture single or more commonly multiple	Benthic, marine, warm shallow water; Permian to Holocene
Suborder Silicoloculinina	Secreted rods of opaline silica		
(1 family, 1 genus)	Imperforate; inner and outer organic membranes, with median layer of randomly arranged siliceous rods bordered above and below by rods that parallel surface	Enrolled, proloculus followed by chambers one-half coil in length, added in varying planes as in miliolines	Benthic, abyssal at high latitudes; Miocene to Holocene
Suborder Lagenina	Optically and ultrastructurally radiate hyaline calcite wall, crystal *c* axis perpendicular to surface, with organic membrane around crystal units		
Robuloidacea (4 families, 30 genera)	Monolamellar, finely perforate; secondary lamination absent or partial in more advanced taxa	Elongate, enrolled, rectilinear, or arcuate; proloculus followed by undivided tubular chamber, or with pseudosepta or true septa and uniserial, or spiraling about long axis; terminal aperture round, slitlike or radiate	Marine, benthic; lower Carboniferous to Lower Cretaceous
Nodosariacea (6 families, 191 genera)	As above, but with secondary lamination resulting from continued growth	Free-living, or rarely attached; single chamber or multilocular, inrolled, arcuate, or rectilinear, or spiraling about long axis; aperture single or multiple, round, slitlike or radiate, and may have entosolenian tube	Marine, benthic; Triassic to Holocene
Suborder Robertinina	Hyaline, of optically and ultrastructurally radial aragonite		
Duostominacea (3 families, 9 genera)	Appears granular; probably was aragonitic but commonly recrystallized; nonlamellar; may include some foreign particles	Planispiral or low to high trochospiral coil, few chambers per whorl; aperture basal, single or double	Marine, benthic; Triassic to Jurassic
Ceratobuliminacea (2 families, 27 genera)	Finely perforate, optically radial aragonite; with primary and secondary lamination, septa primarily bilamellar; surface smooth, pustulose or costate	Trochospiral, few chambers per whorl, chambers with internal partition; aperture basal or areal, may be closed as next chamber is added, and foramen is formed by resorption; may have slitlike peripheral opening	Marine, benthic; Jurassic to Holocene
Conorboidacea (1 family, 3 genera)	As above	Low to high trochospiral coil, later with fewer chambers per whorl to uniserial; aperture basal in trochospiral taxa, terminal in uniserial ones, with hemicylindrical columellalike toothplate	Marine, benthic; Lower Cretaceous to Upper Cretaceous
Robertinacea (1 family, 9 genera)	As above; surface smooth	Low to high trochospiral, transverse inner partition due to infolding of wall produces chamberlets; aperture basal or high loop or both, part remaining as septal foramen; may have supplementary openings	Marine, benthic; Paleocene to Holocene

(*cont.*)

Characteristics of foraminiferan suborders and superfamilies (cont.)

Superfamily	Test wall	Test form	Distribution
Suborder Globigerinina	Perforate hyaline radial calcite wall, *c* axis normal to surface, primarily bilamellar septa		
Heterohelicacea (3 families, 21 genera)	Primary and secondary lamination; surface smooth or striate; may have pore mounds	Trochospiral, triserial or biserial, may reduce to uniserial or show chamber proliferation; aperture a basal arch, terminal in uniserial stage; may have supplementary or accessory openings	Marine, planktonic; Lower Cretaceous to Holocene
Planomalinacea (3 families, 8 genera)	As above, surface smooth, pitted, or with inflational pustules	Planispiral or tending to trochospiral; aperture equatorial, basal, part may remain as relict openings	Marine, planktonic; Lower to Upper Cretaceous
Rotaliporacea (4 families, 17 genera)	As above; surface smooth, pitted, or with inflational pustules; may have imperforate carinal band or keel	Trochospiral; primary basal aperture umbilical-extraumbilical; may also have secondary sutural openings	Marine, planktonic; Lower to Upper Cretaceous
Globotruncanacea (2 families, 17 genera)	Surface smooth, pitted, or with inflational pustules; may have imperforate carinal band	Trochospiral, chambers globular to subangular; primary aperture umbilical, umbilical area with tegilla; may have accessory apertures	Marine, planktonic; Upper Cretaceous
Globorotaliacea (6 families, 40 genera)	Finely lamellar surface smooth, pitted, or cancellate, with inflational pustules or muricae; nonspinose	Trochospiral, chambers rounded to radial elongate; periphery rounded to carinate; primary aperture basal, umbilical-extraumbilical, with imperforate lip; may have supplementary sutural apertures and bullae with accessory openings	Marine, planktonic at medium depths; Paleocene to Holocene
Hantkeninacea (3 families, 8 genera)	Surface smooth to finely pustulate	Planispiral to enrolled biserial; globular to elongate chambers, may have tubulospines; single or multiple equatorial aperture, basal to areal; may have relict secondary openings	Marine, planktonic; Paleocene to Miocene
Globigerinacea (2 families, 24 genera)	Surface smooth or costellate, with nonlamellar, monocrystalline elongate spines set in holes in wall; crystal *c* axes lengthwise of spine	Trochospiral, planispiral, streptospiral, or enveloping; aperture basal and umbilical, umbilical-extraumbilical, or equatorial; may have secondary sutural openings, and may have bullae	Marine, planktonic in surface waters; Eocene to Holocene
Suborder Rotaliina	Perforate hyaline lamellar calcite, deposited at each side of organic lamella; optically radial or granular, according to normal to surface or oblique crystal orientation		
Bolivinacea (2 families, 12 genera)	Optically radial; surface smooth, striate or costate	Biserial, or may later become uniserial; aperture a high loop or slit, basal to areal, with internal toothplate	Benthic, marine; Upper Cretaceous to Holocene
Loxostomatacea (3 families, 7 genera)	Optically radial or granular; surface smooth, pustulose, or costate	Biserial, or may tend to uniserial; aperture single to cribrate, basal to terminal, no toothplate	Benthic, marine; Upper Cretaceous to Holocene
Bolivinitacea (1 family, 2 genera)	Optically radial; surface smooth, striate, reticulate or nodose	Biserial, flattened, truncate margins; one border of high opening bent inward to form toothplate	Benthic, marine; Miocene to Holocene
Cassidulinacea (2 families, 20 genera)	Optically radial or granular; surface smooth, papillate or cancellate	Biserial or derived from such; biserial axis enrolled in planispiral or rarely trochospiral coil, or may uncoil; basal to terminal aperture may have toothplate; may have sutural secondary openings	Benthic, marine; Paleocene to Holocene
Eouvigerinacea (2 families, 6 genera)	Optically radial, surface smooth, hispid or pustulose	Biserial or uniserial, may have early planispiral coil; aperture with hemicylindrical internal toothplate	Benthic, marine; Lower Cretaceous to Eocene; Pliocene
Turrilinacea (3 families, 19 genera)	Optically radial or granular; surface smooth, pustulose, longitudinally costate, striate, or reticulate	Trochospiral or triserial, later may reduce to biserial or uniserial; aperture a basal arch to areal, single to cribrate, may have simple toothplate	Benthic, marine; Jurassic to Holocene
Buliminacea (8 families, 57 genera)	Optically radial; surface smooth, longitudinally costate, hispid, or laterally carinate	High trochospiral or triserial, may reduce to biserial or uniserial; aperture a basal loop to terminal and rounded, internal toothplate attached to apertural margin; rarely with internal subdivisions	Benthic, marine; Upper Cretaceous to Holocene
Fursenkoinacea (2 families, 11 genera)	Optically granular, surface smooth or hispid in early stage	Triserial or twisted biserial; aperture loop-shaped with internal toothplate; may have sutural supplementary openings	Benthic, marine; Upper Cretaceous to Holocene
Delosinacea (3 families, 10 genera)	Optically granular, surface smooth or with sparse basal spines	Trochospiral or triserial, later may be biserial or uniserial; aperture a high basal slit with toothplate; may have sutural pores	Benthic, marine; Lower Cretaceous to Holocene
Pleurostomellacea (1 family, 16 genera)	Optically granular, surface smooth or rarely striate or costate	Triserial or biserial, may become uniserial; straight to curved slit aperture excentric, may have hood, or may be terminal and cribrate; may have internal siphon	Benthic, marine; Lower Cretaceous to Holocene

Characteristics of foraminiferan suborders and superfamilies (cont.)

Superfamily	Test wall	Test form	Distribution
Suborder Rotaliina (cont.)			
Stilostomellacea (1 family, 5 genera)	Optically radial, surface smooth, weakly striate to costate, or with sparsely hispid margins	Uniserial, arcuate to rectilinear; terminal aperture with phialine lip and toothlike projection	Benthic, marine; Upper Cretaceous to Holocene
Annulopatellinacea (1 family, 1 genus)	Optically granular, surface smooth	Conical; proloculus and reniform second chamber, then uniserial with discoid chambers; radial tubules open as pores at surface; plastogamic sexual reproduction	Benthic, shallow marine; Miocene to Holocene
Discorbacea (15 families, 80 genera)	Optically radial, finely to coarsely perforate, imperforate around aperture; surface smooth to pustulose, may have inflational peripheral spines or keel	Low trochospiral coil, rarely uncoils; may have secondary septula; basal aperture on umbilical side or areal and cribrate; may have secondary sutural openings	Benthic, may have planktonic stage; marine; Jurassic to Holocene
Glabratellacea (3 families, 22 genera)	Optically radial, surface smooth to pustulose, may be carinate; umbilical side with radial striae, costae or nodes that aid plastogamy	Low to high trochospire, aperture basal and umbilical; sexual reproduction plastogamic	Benthic, marine; Eocene to Holocene
Siphoninacea (1 family, 5 genera)	Optically radial or granular; commonly with imperforate fimbriate keel	Trochospiral to planispiral, may uncoil, become biserial or uniserial; aperture slitlike to rounded, on neck with phialine lip	Benthic, marine; Upper Cretaceous to Holocene
Discorbinellacea (4 families, 22 genera)	Optically radial, sparsely perforate or lacking pores; surface smooth	Trochospiral, rarely has internal partitions subdividing chamber; aperture basal, an equatorial, low arch, slit up face, or areal; may have supplemental apertures	Benthic, marine; Paleocene to Holocene
Planorbulinacea (6 families, 60 genera)	Optically radial or with intermediate structure, inner and outer layers of calcite crystals normal to the surface and median layer oblique; coarsely perforate	Free or attached; discoidal, lenticular, or conical; trochospirally coiled, uncoiled, rectilinear or biserial, or with chamber proliferation or irregularly added chambers; aperture basal, extraumbilical to equatorial, subterminal if uncoiled; may have supplementary or accessory openings	Benthic, marine; Lower Cretaceous to Holocene
Acervulinacea (2 families, 11 genera)	Optically radial wall; coarsely perforate	Free or attached; may be encrusting, or may grow erect, spiral early stage followed by spreading or irregular chambers; aperture of mural pores or pores in cribrate upper surface of chambers	Benthic, marine warm water; Paleocene to Holocene
Asterigerinacea (7 families, 39 genera)	Optically radial, surface smooth or granulose	Trochospiral to nearly planispiral; chambers with inner partition that attaches to outer wall and forms small umbilical chamberlet; may have distinct equatorial and lateral layers of chambers; aperture basal to areal, on umbilical side to equatorial, may have sutural supplementary openings; may have stolon system	Benthic, marine; Upper Cretaceous to Holocene
Nonionacea (3 families, 30 genera)	Optically granular, rarely optically radial; finely to coarsely perforate	Planispirally coiled to asymmetrical, involute or evolute; aperture simple to cribrate; equatorial, basal to areal, and may have additional peripheral or sutural openings	Benthic, marine; Upper Cretaceous to Holocene
Chilostomellacea (11 families, 69 genera)	Optically granular; finely to coarsely perforate; surface smooth to pustulose	Free or attached; trochospirally coiled, may uncoil or have enveloping chambers; may have inner partition related to toothplate; aperture basal, a vertical slit, umbilical to peripheral, or terminal in uncoiled forms, may have lip orflap, may have secondary sutural openings	Benthic, marine; Lower Cretaceous to Holocene
Orbitoidacea (3 families, 23 genera)	Optically radial, surface smooth or pustulose; may have fissures	Large, discoidal to lenticular, dimorphism prominent, microspheric test coiled or biserial at first, megalospheric test with bilocular embryo; later stage with similar or differentiated equatorial and lateral chambers, intercameral openings modified as stolons, but no canal system	Benthic, marine; Upper Cretaceous to Upper Eocene
Rotaliacea (6 families, 79 genera)	Optically radial, or rarely optically granular, finely to coarsely perforate, surface smooth, pustulose or costate, may have large coarse spines; septal flap attaches to former apertural face as new chamber is added, enclosing system of canals and cavities	Trochospiral or planispiral coiling, evolute to involute, chambers small and numerous, may be subdivided by radial plates or rods; primary aperture single or multiple, basal to areal; sutural openings connect to canal system, not to chamber interior	Benthic, marine; Upper Cretaceous to Holocene
Nummulitacea (4 families, 22 genera)	Optically radial; surface smooth to pustulose; septal flap present as in Rotaliacea	Planispiral, evolute to involute, dimorphism prominent; has many small chambers, may be subdivided into chamberlets, equatorial and lateral chambers may be differentiated, later chambers may be in annular series; aperture single to multiple; may have subsutural canals, spiral marginal cord, spiral canal system, and intraseptal canals	Benthic, marine; Paleocene to Holocene

being replaced by others. The oldest-known and presumably primitive tests are the morphologically simple tectinous and agglutinated ones of the Cambrian and Ordovician, with globular, tubular, branching, or irregular form. By the mid-Ordovician, chambered tests were present (*Reophax*), and in the Silurian, tests with microgranular calcareous walls also had appeared. Many early Paleozoic foraminiferans have been obtained from acid residues or thin sections of rock.

Microgranular tests diversified greatly in the late Paleozoic, true chamber development being followed by many wall and chamber modifications, layered and alveolar walls, labyrinthic chambers, crenulated septa, secondary septa, and chamberlets. Similar wall and chamber modifications were later repeated in other foraminiferal lineages.

Beginning in the Mississippian, various groups of foraminiferans successively were so abundant that their discarded tests accumulated as foraminiferal limestones, particularly noteworthy being the Mississippian endothyrids, Pennsylvanian and Permian fusulines, Cretaceous and Tertiary miliolids, and Eocene nummulites. Although the Fusulinina were abundant and varied in the Paleozoic, they then disappeared whereas the few porcelaneous and hyaline calcareous species that appeared in the late Paleozoic survived to expand greatly in the Mesozoic.

Planktonic foraminiferans became abundant in the Cretaceous and Tertiary; extensive pelagic limestones and chalks represent earlier accumulations like the present *Globigerina* and *Orbulina* oozes on the sea floor.

Ecology. Most foraminiferans are benthic, living upon the sea floor, within the upper few centimeters of ooze, or upon benthic algae or other organisms. They occur from the intertidal zone to oceanic depths, in brackish, normal marine, or hypersaline waters, and from the tropics to the poles. Some modern Lagynacea live in fresh water, but none are known as fossils. Assemblages vary widely in response to local conditions, with the greatest diversity occurring in warm, shallow water. A smaller number, the Globigerinina, are planktonic, living at various depths in the water column from the surface to the bottom, being most numerous between 18 and 90 ft (6 and 30 m). Vertical migration may be diurnal and may occur during ontogenetic development. The preferred depth range of a species may vary geographically in response to temperature differences or to changes in water density.

Abundant distributional data for living species have made possible the extensive use of foraminiferans for paleoecologic interpretation and age dating, their abundance permitting use of statistical methods. Paleoclimatic interpretations are based on temperature-related ratios of right- and left-coiling specimens of trochospiral planktonic species, the latitudinal gradation of species diversity, and the varied oxygen isotope ratios of the test carbonate. Faunal abundance and variety, test size, presence of trace elements, ratios of planktonic to benthic species, and ratios of calcareous to agglutinated tests, as well as gross morphology, have been the basis for other studies of such ecologic parameters as depth of water, distance from shore, sea-level fluctuations, paleocurrents, and relative oxygenation or stagnation of the bottom, reworked or displaced faunas, salinity, pH, turbidity and turbulence, nature of the substrate, rate of detrital influx, and rate of accumulation of sediments. Unsurpassed as stratigraphic markers for petroleum exploration or as tools for paleoecologic interpretation of ancient sediments, foraminiferans have proved to be equally valuable in oceanographic studies of the more recent events in the Earth's history.

Helen Tappan Loeblich

Bibliography. S. Bellemo, *Bull. Geol. Inst. Univ. Upsala*, new series, vol. 6, 1974; H. Bender and C. Hemleben, *Abhandlungen der Geologischen Bundesanstalt*, vol. 41, 1988; R. H. Hedley and C. G. Adams (eds.), *Foraminifera*, vol. 3, 1978; A. R. Loeblich, Jr., and H. Tappan, *Foraminiferal Genera and Their Classification*, 2 vols., 1987.

Force

Force may be briefly described as that influence on a body which causes it to accelerate. In this way, force is defined through Newton's second law of motion. *See* ACCELERATION.

This law states in part that the acceleration of a body is proportional to the resultant force exerted on the body and is inversely proportional to the mass of the body. An alternative procedure is to try to formulate a definition in terms of a standard force, for example, that necessary to stretch a particular spring a certain amount, or the gravitational attraction which the Earth exerts on a standard object. Even so, Newton's second law inextricably links mass and force. *See* MASS.

Many elementary books in physics seem to expect the beginning student to bring to his study the same kind of intuitive notion concerning force which Isaac Newton possessed. One readily thinks of an object's weight, or of pushing it or pulling it, and from this one gains a "feeling" for force. Such intuition, while undeniably helpful, is hardly an adequate foundation for the quantitative science of mechanics.

Newton's dilemma in logic, which did not trouble him greatly, was that, in stating his second law as a relation between certain physical quantities, he presumably needed to begin with their definitions. But he did not actually have definitions of both mass and force which were independent of the second law. The procedure which today seems most free of pitfalls in logic is in fact to use Newton's second law as a defining relation.

First, one supposes length to be defined in terms of the distance between marks on a standard object, or perhaps in terms of the wavelength of a particular spectral line. Time can be supposed similarly related to the period of a standard motion (for example, the rotation of the Earth about the Sun, the oscillations of the balance wheel of a clock, or perhaps a particular vibration of a molecule). Although applying these definitions to actual measurements may be a practical matter requiring some effort, a reasonably logical definition of velocity and acceleration, as the first and second time derivatives of vector displacement, follows readily in principle.

Absolute standards. Having chosen a unit for length and a unit for time, one may then select a standard particle or object. At this juncture one may choose either the absolute or the gravitational approach. In the so-called absolute systems of units, it is said that the standard object has a mass of one unit. Then the second law of Newton defines unit force as that force which gives unit acceleration to the unit mass. Any other mass may in principle be compared with the standard mass (m) by subjecting it to unit force and measuring the acceleration ($\mathbf{a}$), with which it varies inversely. By suitable appeal to experiment, it is possible to conclude that masses are scalar quantities and that forces are vector quantities which may be superimposed or resolved by the rules of vector addition and resolution.

In the absolute scheme, then, Eq. (1) is written

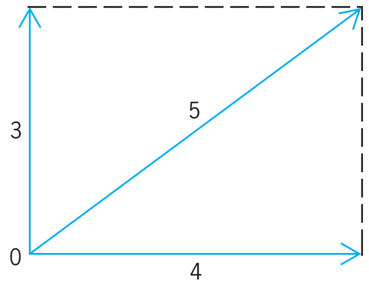
$$\mathbf{F} = m\mathbf{a} \qquad (1)$$

for nonrelativistic mechanics; here boldface type denotes vector quantities. The quantities on the right of Eq. (1) are previously known, and this statement of the second law of Newton is in fact the definition of force. In the absolute system, mass is taken as a fundamental quantity and force is a derived unit of dimensions MLT^{-2} (M = mass, L = length, T = time).

Gravitational standards. The gravitational system of units uses the attraction of the Earth for the standard object as the standard force. Newton's second law still couples force and mass, but since force is here taken as the fundamental quantity, mass becomes the derived factor of proportionality between force and the acceleration it produces. In particular, the standard force (the Earth's attraction for the standard object) produces in free fall what one measures as the gravitational acceleration, a vector quantity proportional to the standard force (weight) for any object. It follows from the use of Newton's second law as a defining relation that the mass of that object is $m = w/g$, with g the magnitude of the gravitational acceleration and w the magnitude of the weight. The derived quantity mass has dimensions FT^2L^{-1}.

Because the gravitational acceleration varies slightly over the surface of the Earth, it may be objected that the force standard will also vary. This may be avoided by specifying a point on the Earth's surface at which the standard object has standard weight. In principle, then, the gravitational system becomes no less absolute than the so-called absolute system. *See* GRAVITATION.

Composition of forces. By experiment one finds that two forces of, for example, 3 units and 4 units acting at right angles to one another at point 0 produce an acceleration of a particular object which is identical to that produced by a single 5-unit force inclined at arccos 0.6 to the 3-unit force, and arccos 0.8 to the 4-unit force (see **illus.**). The laws of vector addition thus apply to the superposition of forces.



Vector addition of forces.

Conversely, a single force may be considered as equivalent to two or more forces whose vector sum equals the single force. In this way one may select the component of a particular force which may be especially relevant to the physical problem. An example is the component of a railroad car's weight along the direction of the track on a hill.

Statics is the branch of mechanics which treats forces in nonaccelerated systems. Hence, the resultant of all forces is zero, and critical problems are the determination of the component forces on the object or its structural parts in static equilibrium. Practical questions concern the ability of structural members to support the forces or tensions. *See* STATICS.

Specially designated forces. If a force is defined for every point of a region and if this so-called vector field is irrotational, the force is designated conservative. Physically, it is shown in the development of mechanics that this property requires that the work done by this force field on a particle traversing a closed path is zero. Mathematically, such a force field can be shown to be expressible as the (conventionally negative) gradient of a scalar function of position V, Eq. (2).

$$\mathbf{F} = -\nabla V \qquad (2)$$

A force which extracts energy irreversibly from a mechanical system is called dissipative, or nonconservative. Familiar examples are frictional forces, including those of air resistance. Dissipative forces are of great practical interest, although they are often very difficult to take into account precisely in phenomena of mechanics.

The force which must be directed toward the center of curvature to cause a particle to move in a curved path is called centripetal force. For example, if one rotates a stone on the end of a string, the force with which the string pulls radially inward on the stone is centripetal force. The reaction to centripetal force (namely, the force of the stone on the string) is called centrifugal force. *See* CENTRIFUGAL FORCE; CENTRIPETAL FORCE.

Methods of measuring forces. Direct force measurements in mechanics usually reduce ultimately to a weight comparison. Even when the elastic distortion of a spring or of a torsion fiber is used, the calibration of the elastic property will often be through a balance which compares the pull of the spring with a calibrated weight or the torsion of the fiber with a torque arising from a calibrated weight on a moment arm. *See* BALANCE.

In dynamic systems, any means of measuring acceleration—for example, through photographic methods or radar tracking—allows one to calculate the force acting on an object of known mass.

Units of force. In addition to use of the absolute or the gravitational approach, one must contend with two sets of standard objects and lengths, the British and the metric standards. All systems use the second as the unit of time. In the metric absolute system, the units of force are the newton and the dyne. The newton, the unit of force in the International System (SI), is that force which, when applied to a body having a mass of 1 kilogram, gives it an acceleration of 1 m/s^2. The poundal is the force unit in the British absolute system, whereas the British gravitational system uses the pound. Metric gravitational systems are rarely used. Occasionally one encounters terms such as gram-force or kilogram-force, but no corresponding mass unit has been named. *See* UNITS OF MEASUREMENT.

George E. Pake

Bibliography. F. Bueche and D. A. Jerde, *Principles of Physics*, 6th ed., 1995; D. Halliday, F. Resnick, and K. Krane, *Physics*, 4th ed., 1992; F. J. Keller, W. E. Gettys, and M. J. Skove, *Physics*, 2d ed., 1993; C. Kittel, W. D. Knight, and M. A. Ruderman, *Berkeley Physics Course*, vol. 1: *Mechanics*, 2d ed., 1973.

Force analysis

An analysis yielding the respective forces acting at any point of any member, or part of a member, of a mechanism, obtained by using relationships for dynamic equilibrium in a plane rigid body subject to external forces within this plane and to internal forces due to its motion in this plane. The following treatment is a careful selection and condensation of essential features of the reasoning.

The basic relationships for equilibrium are that the forces acting upon a body, or their resultants, must act through a point (their force polygon is closed); and that the couples acting on the body also reduce to zero (the moment of the existing forces on any member about any point is zero). If there are no more than three unknowns, a solution can be found, utilizing the three conventional equations summing forces in two perpendicular directions to zero and summing moments to zero. If there are more unknowns, further information must be acquired from other members.

The method of approach may, in practice, be purely algebraic, for instance, using the three mentioned equations of force and moment equilibrium; or it may be graphical, using diagrams adapted to yielding quantitative results. It is often advantageous to combine the two methods where some of the quantities used graphically are cumbersome and can be more easily computed.

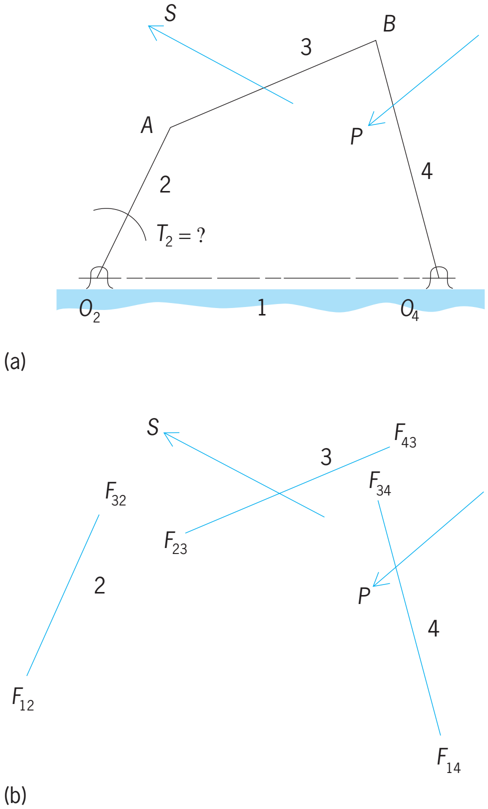


Fig. 1. Construction of free-body diagram. (*a*) Forces *S* and *P* acting on members 3 and 4 of a four-bar linkage. (*b*) Free-body diagrams of the three moving members of *a*.

Free-body diagram. A fundamental tool in solving problems is the free-body diagram. In this, each part of the mechanism is isolated and the forces acting upon each part are established. It is then often possible to piece this knowledge together until the entire solution is known. In **Fig. 1***a*, a four-bar linkage with frictionless pins and weightless members is acted upon by two forces *S* and *P* on members 3 and 4, respectively. The forces on all pins and the couple on member 2 are sought. Figure 1*b* shows the three moving members isolated in their respective free-body diagrams. Using these diagrams, if one starts with 4, one can successfully work backward for the

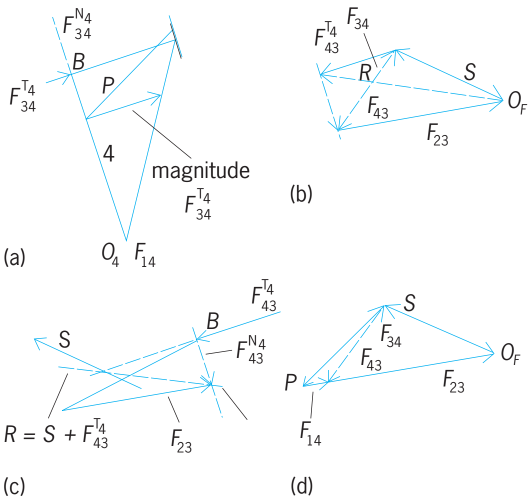


Fig. 2. Successive steps (*a-d*) in solution of forces.

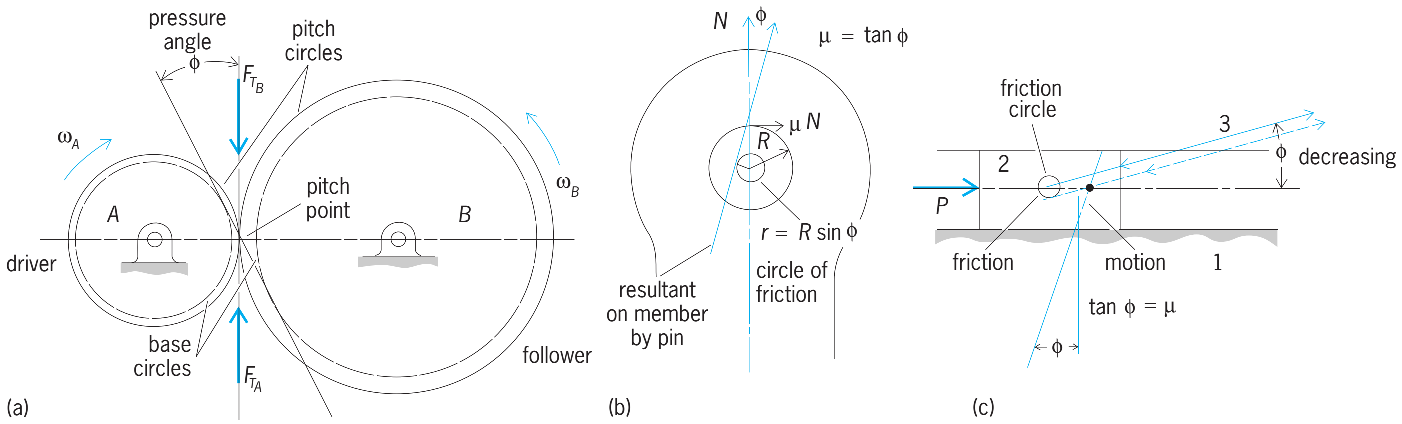


Fig. 3. Force diagrams. (*a*) Force between two gears. (*b*) Force on a pivot or pin of radius *R*. (*c*) Force on the upper or lower side of a slider.

entire solution. In **Fig. 2***a*, a moment diagram with fulcrum at O_4 permits solution for F_{34}^{T4}, the normal force component of member 3 on member 4 at pin *B*. In Fig. 2*b* (F_{34}^{T4}, *S*, and *R* are now known), F_{23} and F_{43}^{N4} can be found. In Fig. 2*c*, knowledge of F_{43}^{N4} yields the direction of *R*. Figure 2*d* thus reveals the direction and magnitude of F_{14} on pin O_4, as well as the pin force (equal and opposite to F_{23}), and hence the couple on 2. *See* FOUR-BAR LINKAGE.

Force diagram. Force diagrams for many friction and other devices common to mechanisms may be helpful. The force between two gears is one through the pitch point in the direction of the common tangent to the two base circles and inclined at the pressure angle to the common tangent to the pitch circles (**Fig. 3***a*). The force acting on a pivot or pin of radius *R* will act along the center line and through the center point unless there is friction in the pin (Fig. 3*b*), in which case the force will lie at the friction angle where the center line meets the moving head, at angle ϕ ($\tan \phi = \mu$). Hence it is tangent to a small circle, the friction circle, $r = R \sin \phi$ (Fig. 3*b*). The force on the upper or lower side of a slider will vary in direction with forward and return strokes, but will lie at the friction angle to the normal and through the point of concurrency of other forces (Fig. 3*c*). *See* GEAR.

In **Fig. 4**, a press mechanism with massless members but with pin and slider friction is treated. There is a known resistance force of *P* units at output slider

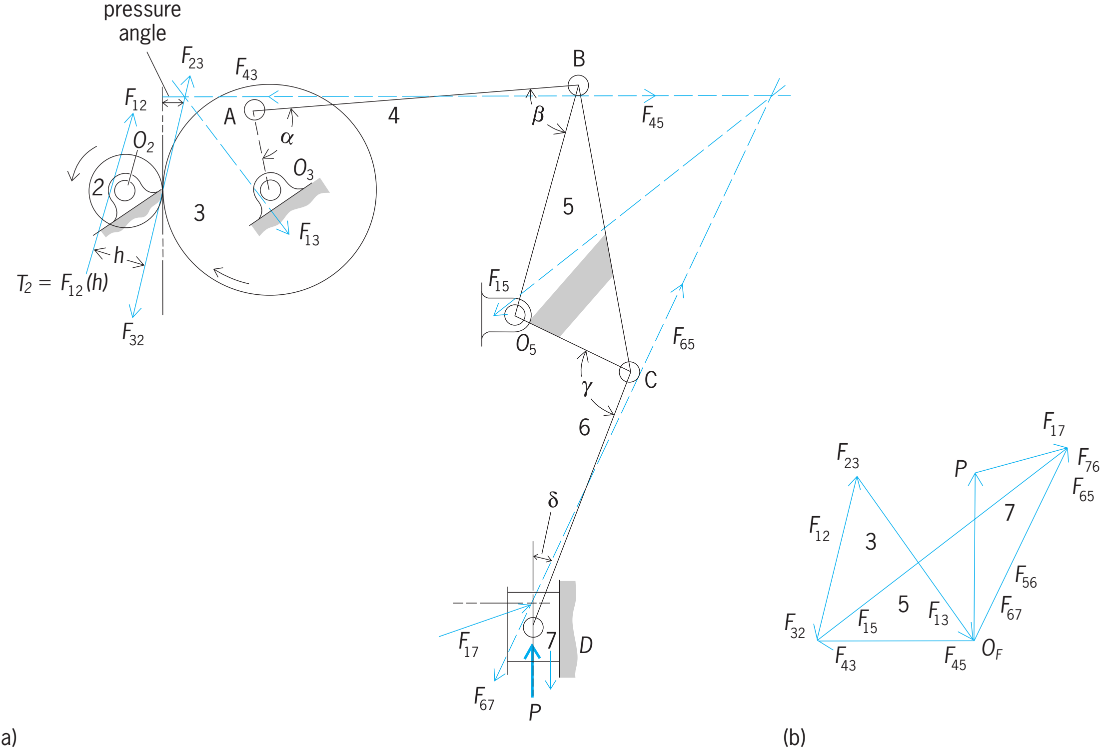


Fig. 4. Solution of (*a*) force problem by (*b*) force triangles.

block 7. The solution polygon showing all forces appears in Fig. 4*b*. All joint frictions have been assumed equal, with gear friction not being treated. The general method of solution in Fig. 4*a* observes that polygon triangles can be evaluated backward, starting with the three forces acting on 7. Respective part directions must first be ascertained followed by general force directions under impending motion. Next, in order to fix the placement and direction of friction forces, it is necessary to decide whether the angles are respectively increasing or decreasing. This can be done by inspection, by diagrams, or by computations. With the direction of each of the forces at the pins known, it can be determined upon which side of the pin the force must pass to induce a rotation opposite to that due to joint friction. Thus at point A, the geometry shows that angle α is expanding; hence friction must be acting clockwise (to oppose it), and hence F_{43} must act on the upper side of the friction circle to turn it counterclockwise, as shown. Corresponding patterns apply for angles β, γ, and δ. Thus at D, since δ is diminishing under the impending motion, pin motion is counterclockwise, pin friction is clockwise, and the force F_{67} must yield a counterclockwise effect. At D there is a backward, upward friction force on the left side of the slider, establishing the friction force.

Once the positions of the forces are known, as in Fig. 4*a*, three basic force triangles are successively solved; in each, the direction and magnitude of one force and the directions of two others are known. Thus with P known, F_{65} is ascertained for the triangle of forces involved with body 7. For the force triangle for body 5, F_{56} has just been found, and so F_{45} can be found in direction and magnitude. For the force triangle for body 3, F_{45} is known in position and magnitude. Hence, with knowledge of the directions of F_{32} and F_{13}, the complete triangle is known, and F_{12} can be ascertained in direction and magnitude.

When inertia forces and moments are involved, as occurs in moving parts due to their masses, the situation can be met by regarding the forces momentarily as static forces in equilibrium. By using such devices, the existing forces (such as compression forces or shear forces) and couples on mechanisms can frequently be found at points of interest. It is assumed that this mass is concentrated at the known center of gravity and that its linear and angular accelerations have been learned from kinematic studies. Kinematics is the inevitable background ingredient of every preparation for a force study where mass is present.

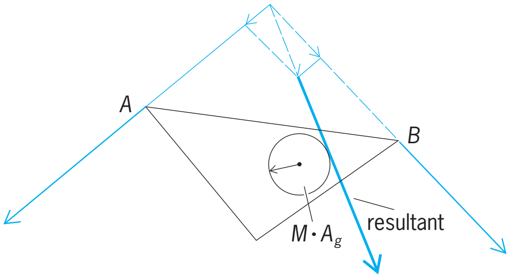


Fig. 5. Representation of the forces acting on a mass to cause the resultant force.

The force $F = MA_g$ is thus known in direction and magnitude, and Eq. (1) holds. Here I is the mass

$$T = I\alpha \qquad (1)$$

moment of inertia (the inertia of rotation) and T is the turning force, torque, or inertia moment. Then Eq. (2), or Eq. (3), can be written. Here k is the radius of gyration.

$$MA_gH = I\alpha \qquad (2)$$

$$h = \frac{I\alpha}{MA_g} = \frac{k^2\alpha}{A_g} \qquad (3)$$

Equations (1)–(3) can also be shown in circle form (Fig. 4*a*), where the force acting upon the mass (to produce the known acceleration) is represented as the resultant of the other forces upon the piece (**Fig. 5**). *See* MOMENT OF INERTIA. Douglas P. Adams

Bibliography. E. A. Dijksman, *Motion Geometry of Mechanism*, 1976; H. H. Mabie and F. W. Ocvirk, *Mechanisms and Dynamics of Machinery*, 4th ed., 1987; B. Paul, *Kinematics and Dynamics of Planar Machinery*, 1979.

Force fit

A means for holding mating mechanical parts in fixed position relative to each other. In the case of a force fit of cylindrical parts, the inner member has a greater diameter than the hole of the outer member; that is, the metals of the two parts interfere. *See* ALLOWANCE.

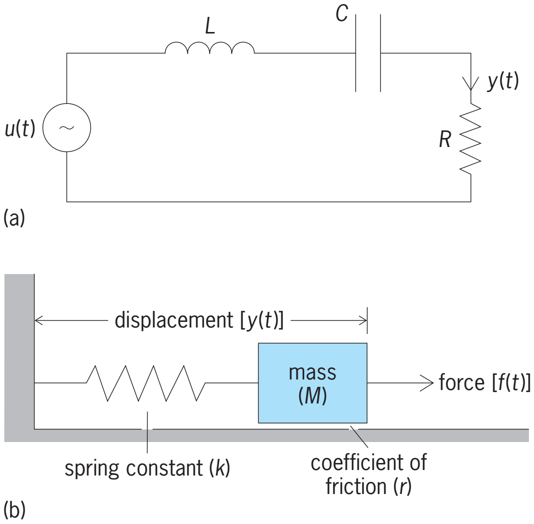
Light drive fits require assembly pressure and produce a more or less permanent assembly of thin sections or long fits. They are suitable for cast iron external members. For a tighter fit, medium or heavy drive fits are used, although such fits may require more than direct mechanical pressure for assembly. In a true force fit, the parts are highly stressed, the interference amounting to 0.002 or 0.003 in. (0.0051 or 0.0076 cm) for parts with a basic diameter of 1 in. (2.5 cm). *See* SHRINK FIT. Paul H. Black

Forced oscillation

A response of a mechanical or electrical system in reaction to an external signal.

Electrical systems. A simple *RLC* circuit (**illus.** *a*) consists of a resistor with resistance R (measured in ohms), an inductor with inductance L (measured in henrys), and a capacitor with capacitance C (measured in farads). The dynamics relating the input voltage, $u(t)$, to the current, $y(t)$, passing through the resistor are described by Eq. (1). Equation (1)

$$L\frac{dy}{dt} + Ry + \frac{1}{C}\int y\,dt = u(t) \qquad (1)$$



Examples of forced oscillation. (*a*) Electrical system, an *RLC* circuit. (*b*) Analogous mechanical system, a spring-mass-damper system.

states that the input voltage is equal to the sum of the voltage across the inductor, the voltage across the resistor, and the voltage across the capacitor, where the voltage across the inductor is the product of its inductance (L) and the rate of change of the current through the inductor; the voltage across the resistor is the product of its resistance (R) and the current passing through it; and the voltage across the capacitor is the integral over time of the current through the capacitor (that is, the charge on the capacitor plates) divided by the capacitance (C).

A fundamental property of differential equations states that the response of a differential equation to a periodic input can be decomposed as a sum of two responses. The first one, called the zero-input response or free oscillation, is due to initial energy stored in the circuit and decays eventually to zero. The second one, due to the voltage input $u(t)$, converges to a periodic signal with the same frequency as $u(t)$. The latter is referred to as the forced oscillation or the steady-state response. The decaying rate of the free oscillation depends on the time constant of the circuit which is determined by the values of R, L, and C and the structure of the circuit. *See* TIME CONSTANT.

Impedance. If the input voltage $u(t)$ is a sinusoidal function with magnitude A (measured in volts) and angular frequency ω (measured in rad/s), then the forced oscillation of $y(t)$ is also a sinusoidal function with the same angular frequency, whose magnitude (B) and phase (θ) with respect to $u(t)$ depend on a quantity called the complex impedance (Z). The complex impedance is a complex number whose real part is just the resistance (R) and whose imaginary part equals $\omega L - 1/(\omega C)$. The magnitude of the forced oscillation (B) is equal to the magnitude of the input voltage (A) divided by the modulus of the complex impedance ($|Z|$). The phase (θ) of the forced oscillation with respect to the input voltage is just the negative of the argument of the complex impedance. Since the complex impedance is a complex number, its modulus is the square root of the sum of the squares of its real and imaginary parts, and its argument is the inverse tangent of the ratio of its imaginary part to its real part. *See* ALTERNATING-CURRENT CIRCUIT THEORY; COMPLEX NUMBERS AND COMPLEX VARIABLES; ELECTRICAL IMPEDANCE; PHASE (PERIODIC PHENOMENA).

Damping. The free oscillation is independent of the input voltage $u(t)$. It can take the form either of an exponentially decaying function or of an exponentially decaying sinusoidal function, depending on the damping coefficient, $R/(2L)$, and the undamped resonant frequency, $1/\sqrt{LC}$. If the damping coefficient is greater than the undamped resonant frequency, then the free oscillation takes the form of an exponentially decaying function; and if the damping coefficient is less than the undamped resonant frequency, then the free oscillation assumes an exponentially decaying sinusoidal function. In contrast, the forced oscillation is independent of the damping coefficient and the undamped resonant frequency. It always has the same frequency as the input signal $u(t)$. *See* DAMPING.

Mechanical systems. Similarly, an analogous mechanical system, a simple spring-mass-damper system (illus. *b*), consists of a body with mass M, which is attached to a wall by a spring with spring constant k, and rests on a horizontal surface over which it moves with friction coefficient r. The dynamic equation that relates the force applied to the body, $f(t)$, to the body's displacement, $y(t)$, is given by Eq. (2).

$$M\frac{d^2y(t)}{dt^2} + r\frac{dy(t)}{dt} + ky(t) = f(t) \qquad (2)$$

Equation (2) states that the force applied to the body equals the sum of the three quantities: the product of the body's mass and its acceleration, the negative of the frictional force, and the negative of the force exerted by the spring. Here, the negative of the frictional force is the product of the coefficient of friction and the body's velocity, and the negative of the force exerted by the spring is the product of the spring constant and the body's displacement. Moreover, the body's velocity is the first derivative of its displacement with respect to time, and its acceleration is the second derivative of its displacement with respect to time.

Analogous to the *RLC* circuit case, application of a sinusoidal force $f(t)$ results eventually in a forced oscillation of the displacement $y(t)$ that is also a sinusoidal function. The magnitude and the phase of the displacement $y(t)$ depends on the complex mechanical impedance that is a function of the mass (M), the spring constant (k), and the friction coefficient (r). The exact evaluation is similar to the *RLC* circuit case. *See* MECHANICAL IMPEDANCE; MECHANICAL VIBRATION; VIBRATION.

Resonance. In both examples, the magnitude of the forced oscillation depends on the complex impedance Z. The imaginary part of Z will be zero if ωL equals $1/(\omega C)$. The value of the angular frequency

ω that satisfies this equation is given by Eq. (3). At

$$\omega_o = \frac{1}{\sqrt{LC}} \qquad (3)$$

this value of ω, the impedance becomes a real number and the maximum magnitude of the oscillation occurs. The frequency ω_o is called the resonant frequency, and the system at this frequency is said to be in resonance. *See* RESONANCE (ACOUSTICS AND MECHANICS); RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Quality factor *Q*. A closely related concept is the quality factor *Q*. In the context of the *RLC* circuit, *Q* is defined by Eq. (4). It can be shown that the ratio of

$$Q = \frac{\omega_o L}{R} = \frac{1}{\omega_o RC} \qquad (4)$$

the magnitude of the output voltage $y(t) \cdot R$ to that of $u(t)$ is given by Eq. (5). This equation shows that

$$\frac{BR}{A} = \frac{1}{\sqrt{1 + Q^2\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right)^2}} \qquad (5)$$

if *Q* is large, the ratio is small unless ω is close to ω_o. In other words, high *Q* corresponds to narrow bandwidth and low *Q* corresponds to wide bandwidth.

The quality factor *Q* also has a different interpretation, which can be explored via an energy analysis. In fact, *Q* is 2π times the ratio of the maximum energy stored in the system to the energy dissipated per cycle. There is a continuous exchange of energy between the magnetic field of the inductor and the electric field of the capacitor. During each half cycle, the energy stored in the capacitor's electric field varies from a maximum value to zero and back to the maximum value again. The energy stored in the inductor's magnetic field varies in a similar manner. The way in which energy exchange takes place is as follows: During one-quarter cycle the capacitor takes in energy as it is released by the inductor, and during the following one-quarter cycle the inductor absorbs energy as the capacitor discharges it. The total energy stored in the circuit is a constant, though the energy stored in either capacitor or inductor varies with time. Similarly, in the mass-spring system, there is a continuous exchange between the kinetic energy of the mass and the potential energy of the spring, but the total energy of the system is constant. *See* HARMONIC MOTION; HARMONIC OSCILLATOR; OSCILLATION.

Er-Wei Bai

Bibliography. R. C. Hibbeler, *Engineering Mechanics*: *Dynamics*, 8th ed., 1997; J. D. Irwin and R. M. Nelms, *Basic Engineering Circuit Analysis*, 8th ed., 2005.

Forensic anthropology

The application of physical anthropology theory and techniques to answering questions for the law. Most human remains are soon discovered and thus available for analysis. Medical evidence and biological techniques used on the living can be applied to these intact remains, as can extensive internal examinations and histological preparations. However, burned, decomposed, mutilated, or fragmented bodies contain less soft tissue evidence and are therefore more difficult for the medical examiner or forensic pathologist to analyze. In these cases, the information provided by the skeleton becomes more important, and may yield the only information available about the identity of the individual and the circumstances surrounding death. Age, sex, ancestry, stature, muscularity, handedness, habits, occupational activities, disease, and injuries can be ascertained from elements of the skeleton, and in such cases the services of a forensic anthropologist may be a valuable adjunct to the work of the forensic pathologist or coroner.

Anthropology is traditionally divided into three major areas: cultural anthropology, which deals with the social study of human cultures; archeology, which deals with past human societies; and physical (or biological) anthropology, which concentrates on humans and their primate relatives. Physical anthropology includes the variation found in the anatomy of the human body and how the body has evolved through time, how it varies in existing populations around the world, how it adapts to different environments, and how it changes during life. Not all physical anthropologists are specialists in the human skeleton, and in fact, even those who are human osteologists may not be experienced in all aspects of skeletal identification and analysis. Training is required in such areas as skeletal variation in current populations or injuries from modern weapons. Many physical anthropologists, however, are increasingly involved in the forensic applications of their expertise, a few on a full-time basis. Recently, archeologists have been asked to aid in the excavation of human remains from clandestine graves, and so the field of forensic archeology has arisen, but the term "forensic anthropology" usually refers to the biological aspects. *See* ANTHROPOLOGY; ARCHEOLOGY; PHYSICAL ANTHROPOLOGY.

In the field. Qualified forensic anthropologists must be actively involved at the scene of the incident to locate, document, map, and recover scattered, buried, decomposed, burned, or dismembered remains. Accurate documentation of the location and condition of remains is important in the identification of the remains and in the reconstruction of the incident. Whether in death investigations involving one body or in a mass fatality incident, the investigator must always assume that litigation will occur and must be prepared to document every move in preparation for a trial. Anthropologists with extensive training in archeological location and documentation techniques (sometimes with forensic archeologists) are extremely valuable at a crime or other recovery scene, and often discover and document materials and context (the relative location of materials to each other within a scene) overlooked by untrained law enforcement personnel. Recovery of crime scene materials should always be performed with a crime scene technician, however.

Human versus nonhuman. This determination is usually simple for the forensic anthropologist but is critical. An investigation can begin or end with this decision. A wide variety of nonhuman remains have been mistakenly identified as human, including burned rabbits, bear paws and other bones, goats, turtle shells (mistaken as skull fragments), and chicken bones.

Identification of scattered remains. The ability to recognize human remains as they are found can assist in many immediate decisions concerning the location and nature of additional remains, and aid in the identification of the circumstances surrounding death. Since the more than 200 bones from an adult skeleton may be scattered over several acres, it is necessary for the forensic anthropologist to determine what bones have been found and what portions are still missing. The relative locations of various skeletal elements and associated evidence give clues as to the site of initial deposition of the body and the agents responsible for scattering the remains.

Recovery of buried remains. It is certainly possible to train crime scene investigators or other personnel to recognize intact human bones and to conduct proper search and excavation of human remains. Still, the forensic anthropologist may be able to offer much better documentation of the exact location of objects in the burial environment, as well as to better interpret the date and shape of the initial excavation at the time of the burial of the body. Many anthropologists or archeologists have the highly developed skills of excavation that can even determine the tool used to dig the original grave. In complex situations, a forensic anthropologist, archeologist, and trained crime scene team are recommended.

Burned bodies and mass fatality incidents. Scenes such as airplane crashes and building explosions and fires can be complicated and may overwhelm the local authorities and facilities. Forensic anthropologists are increasingly members of teams of identification specialists (such as odontologists, pathologists, fingerprint experts, and radiologists) called to help identify victims of mass fatality incidents. Recognition of bones becomes more difficult in these instances, because when burning has destroyed the soft tissue, the bones may be black, gray, or even white. Also, burned bone is extremely fragile and can disintegrate with little movement, and the burning process breaks, warps, and shrinks bones, sometimes up to 20%.

In the morgue or laboratory. It is important to involve forensic anthropologists, odontologists, pathologists, and radiologists as a team working together to identify remains. If the anthropologist is allowed to see only those remains unidentified by the other team members, it will be more difficult to reassociate body parts to identified bodies. An initial examination of the remains is ideally conducted jointly by the forensic pathologist and forensic anthropologist. As sharp instruments are used by the forensic pathologist in the autopsy process, the anthropologist is often present to document this additional damage to the remains (to avoid confusion in the analysis of perimortem trauma later). At this point, evidence (such as insect remains or degree and type of decomposition) for the determination of time since death is collected or observed before forensic examination causes further alteration of the remains. Maceration is the removal of all soft tissue from the bones. Usually the forensic anthropology laboratory will macerate the remains if necessary.

Minimum number of individuals. When incomplete skeletons are found, the number of individuals represented must be determined with care by checking duplication of skeletal elements, symmetry between sides of paired bones, and the size of bony articulations. Commingling is always considered a possibility.

Determination of sex. The sex of the remains can be determined by the morphology of many skeletal elements, particularly the pelvis. The expert can determine sex from more or less complete skeletons with a high degree of reliability, except in the case of subadult skeletons. Overall size and markings of muscle attachments are useful in almost all bones. Measurements and statistical tests can also be used to support conclusions.

Determination of age. Age can be determined from the human skeleton by a great many indicators. Which indicators are appropriate depends upon the general age of the deceased and the condition of the remains. A very young individual normally shows a wealth of maturation information associated with growth and development, including dental development and eruption, and appearance of centers of bone growth. Differences in skeletal age and dental age are frequently observed, particularly in females. The age estimates are usually precise within a small range. During the middle to late teen years, growth nears completion and age estimates may have more variability. A few areas of skeletal maturation often persist until the beginning years of the fourth decade (30 to 32). Skeletal remains of the mature adult still exhibit changes related to age, but most are probably degenerative in nature. Age estimates based upon maturation are usually much more precise than those derived from degenerative changes.

Determination of ancestry. The cranial and facial features are the best biological indicators of ancestry. These may be demonstrated by statistical techniques and metric data from the cranium. However, determination of ancestry is one of the most difficult diagnoses to make in forensic anthropology, in part because there is more variation within a race than between races. In fact, one cannot take one or two traits and assign every skeleton in the world to one of the commonly defined races. Moreover, while the forensic anthropologist is diagnosing ancestry by using biological features, an individual is classified (or classifies himself or herself) to an ancestral or ethnic group based on social as well as biological features.

Stature estimation. Measurements of individual long bones are the best source for stature estimates.

Formulas specific for males and females and various ancestries are available for each of the six major long bones. When very complete skeletons are available, individual measurements of all bones that contribute to stature can yield good estimates with several techniques. Measurements of incomplete long bones can also yield stature estimates, but they are obviously less precise.

Idiosyncrasies. Occasionally, the skeleton shows unusual variants, such as a sternal foramen, nonclosed posterior arches of the sacrum, bifurcated ribs, or cervical or lumbar ribs. Very often, asymmetry of the face, limbs, or jaws will be present. These may aid in the identification process.

Disease, trauma, and time of death. Many diseases and injuries suffered by the individual during life will leave evidence in the skeletal system. These include arthritis, bone infections (including tuberculosis, syphilis, and osteomyelitis), cancers, developmental disorders, healed fractures, orthopedic surgical procedures, or surgically implanted devices that might be embedded in the bones or surrounding soft tissue. The active diseases present in the victim around the time of death, or injuries received at the time, are always important considerations in any death investigation. Conclusions about the cause and manner of death must be made by the pathologist, or in some jurisdictions, the coroner; but it is the role of the forensic anthropologist to describe and interpret any evidence present in the skeletal remains that may be useful for the pathologist to reach informed conclusions. Without clinical experience, the anthropologist normally would not offer testimony or opinions concerning amount of pain to the victim, how long consciousness would have been maintained, or whether an injury would have been lethal. Some cases may be exceptional, such as a shotgun wound resulting in massive destruction of the cranial vault. Likewise, injuries to the now absent soft tissue may be inferred by the damage to the bones. For instance, the wounds which caused cut marks from the sharp edge of a weapon on the anterior surfaces of the bodies of the cervical vertebra would have likely damaged the vascular structures in the neck.

Forensic anthropologists can often interpret postmortem damage to the remains and use this information to help diagnose the elapsed time following death by taking into consideration the environmental conditions (including climate, depth of deposition, and coverings over the body) and the degree of decomposition, skeletonization, scavenging, and weathering of the remains. Very fresh bodies can yield precise estimates of the time of death, but as time passes, less precision can be achieved. Estimates of time since death for decomposed remains are broad, but they reflect the variety of reactions that the remains can have to environmental conditions. Narrow estimates will produce more difficulties in the investigation or prosecution than broad estimates.

Facial approximation. Publication of photographs of three-dimensional sculpting over the skull and two-dimensional drawings based on photographs of the skull is a way of seeking information from the public about the identity of victims. This technique is a last attempt to identify a victim, and is used when the investigators do not know of any possible victims, or when all known possible victims have proven to be incorrect. The names proposed by the public as a result of the reconstructions can then be checked by the investigators. If one appears to be correct after further investigation, then one of the above techniques should be used to confirm the identification. Facial reconstructions do not confirm identifications, but they are means of seeking possible victims. Many forensic anthropologists do not do facial reconstructions.

Photo superimposition. This relatively new technological application is used to compare skulls with antemortem photographs of possible victims. The technique usually employs two video cameras, one imaging the skull and the other the photograph of the person when alive. A modified version of this means of narrowing the identities of individuals involves the use of photographs taken at automated teller machines, trap camera, or bank cameras compared with photographs taken of living suspects or victims.

Positive identification. Final identification of the remains as those of a particular person may be made by a number of means. The most readily accepted means of positive identification ("to the exclusion of all others") include dental identification (comparing the dental records before and after death), fingerprint analysis, and DNA. Forensic anthropologists often aid in the collection of evidence for these means of identification, but most forensic anthropologists are not qualified to make the final identifications. If enough individualizing characteristics are discovered in skeletal elements (through radiographic or gross comparison of information before and after death), positive identification may be made in court from those alone. *See* ANTHROPOMETRY; FORENSIC CHEMISTRY; FORENSIC MEDICINE. Diane L. France

Forensic biology

The scientific analysis of biological evidence to provide objective information on legal matters or those that pertain to criminal and civil law. Biological evidence such as bodily fluids or tissues that may be found at crime scenes can be analyzed through deoxyribonucleic acid (DNA) typing. Typing requires detection and screening of the biological evidence, such as blood, semen, or saliva, extracting the DNA from a specimen, amplifying specific regions of the DNA using the polymerase chain reaction (PCR), and typing the resulting PCR products to determine a DNA profile. The DNA profile from the evidence is then compared to known profiles from suspects, victims, or database samples to determine the significance of the result. Samples containing mixtures require additional interpretation to infer individual donor allele designations. Forensic biologists must also assess the statistical significance of their results, write reports, and testify in court. *See* CRIMINALISTICS; DEOXYRIBONUCLEIC ACID (DNA); FORENSIC EVIDENCE.

The establishment of a United States national DNA database, the Combined DNA Index System or CODIS, has facilitated the ability to compare DNA profiles from unknown biological crime scene evidence to DNA databases of known convicted criminals (“cold hits”) or to DNA left at other crime scenes, resulting in the ability to link cases. As of 2005, CODIS contained over 2.5 million profiles. Many other countries, such as the United Kingdom and Germany, have DNA databases with some of the same genetic markers, permitting international database searches. In the United States alone, over 25,000 investigations have been aided through CODIS.

By comparison of the DNA profile from crime scene samples to known samples, the results can serve to link victims and suspects with the crime scene or can exclude a suspect from association with that crime. Additionally, scientific analysis of biological evidence may provide unbiased information to substantiate case circumstances, corroborate or refute an alibi, or identify a weapon used in a crime. Cases may include nonhuman samples such as botanical, fungal, entomological, or zoological specimens, which can also be used to link victims and suspects with each other or to the crime scene. Forensic biology may also be applied to homeland security, in which crimes may include weapons of mass destruction such as pathological microbes. The forensic biologist may be able not only to identify the microbe (such as anthrax) but to type the microbial DNA and link the strain of the microbe to a strain produced at a certain location or to the original progenitor strain. *See* FORENSIC BOTANY; FORENSIC ENTOMOLOGY; MICROBIOLOGY; PATHOLOGY.

Two main principles permit the use of DNA in forensics. First, no two individuals have the same DNA, with the exception of identical twins. Second, the DNA from any source (such as blood, hair, or skin) of a particular individual will be the same.

Evidence detection and screening. Detection starts with evaluating the investigative information to understand the nature of the case and samples. Visual examination of stains with alternative light sources (such as ultraviolet or infrared light and lasers) and chemical enhancement reagents may be performed. Forensic biologists can determine the nature of the biological stain (for example, blood, semen, or saliva), and whether it is human through presumptive and confirmatory tests. These tests consist of analytical procedures including microscopy, chemical tests, and immunological assays. *See* FORENSIC CHEMISTRY; FORENSIC MICROSCOPY.

DNA extraction. Following the detection and screening of samples from a crime scene, DNA must be extracted. Among the methods used are the following:

1. *Organic extraction*. This method consists of lysis of the cells in a detergent-based buffer followed by one or more rounds of purification using an organic phase separation and concentration using column centrifugation or ethanol precipitation.

2. *Chelex resin extraction*. This method utilizes a fast, simple extraction of small amounts of sample in the presence of a chelating resin. It results in a somewhat crude extract but is usually adequate for PCR amplification of the forensic genetic loci.

3. *Solid-phase extraction*. These methods utilize a membrane that act as a capture device for the DNA. Samples are spotted onto the membranes, and the subsequent washes remove the impurities.

4. *Silica-based extraction*. Nucleic acids are first adsorbed to the silica in the presence of chaotropic salts (salts that weaken hydrophobic associations). Polysaccharides and proteins do not adsorb and are removed. Next, following a wash in low salt, pure nucleic acids are released. This method has been automated using robotic stations and is being used in several laboratories. *See* EXTRACTION.

DNA quantification. Assessing the quantity and quality of the sample is the next step. Several methods are being utilized in crime laboratories. These include agarose gel electrophoresis in the presence of quantification standards (samples with known quantities of DNA), known as yield gel electrophoresis; slot blot hybridization, using known DNA standards immobilized on a membrane followed by hybridization to a human/higher primate–specific DNA probe; homogeneous plate assays, using a DNA fluorescent dye and scanning in a plate reader; and more recently, real-time detection using quantitative PCR (QPCR). Real-time QPCR has several advantages over the other methods in that it is extremely accurate and sensitive over a broad dynamic range, and it occurs in a closed-tube system, reducing the potential for carryover contamination. Using this technique, a forensic biologist can monitor and quantify the accumulation of PCR products during log phase amplification.

DNA amplification. PCR is a fast in-vitro DNA synthesis process that can provide up to a billion (10^9) copies of a given target sequence. Specific DNA markers can be targeted for duplication by a DNA polymerase. Primers are designed to hybridize to the specific markers along the length of the DNA template during the cycling of temperatures. In the thermal cycle, DNA strands are separated, primers bind to the template, and then a special DNA polymerase that is heat-stable is used to copy and amplify the genetic markers. Through a process of 28–32 heating and cooling cycles, the DNA is increased so that it can be analyzed. The thermal cyclers contain many sample wells, permitting the amplification of multiple samples simultaneously; as many as 96 samples may be amplified in under 3 hours. *See* FORENSIC MEDICINE; GENE; GENE AMPLIFICATION; MOLECULAR BIOLOGY.

Multiplex PCR allows several different loci to be simultaneously amplified in a single tube. This enables typing from a single aliquot of the extracted genomic DNA, reducing sample consumption. Recently the ability to analyze as many as 15 autosomal short tandem repeats (STRs) simultaneously has been reported using DNA from a very small amount of degraded sample.

DNA separation and detection. Separation and detection of the amplified products is required following PCR. There are many different methods to achieve typing. These include polyacrylamide gel electrophoresis (PAGE) followed by silver staining or, if the primers are fluorescently tagged, detection by fluorescent gel scanners; and capillary electrophoresis (CE) with laser-induced fluorescence. The latter method has become the most commonly utilized method of detection, as it is highly automated (there is no gel to pour and load), samples can be easily be reinjected (robotically), and since the DNA traverses the entire length of the capillary, the resolution of the higher-molecular-weight loci is usually better than in the PAGE methods. *See* ELECTROPHORESIS.

Forensic genetic markers. The genetic markers currently being typed in most forensic biology laboratories include autosomal short tandem repeats (STRs), mitochondrial DNA, and Y chromosome STRs. *See* GENETIC MAPPING.

Short tandem repeats. Short tandem repeats consist of regions of 2–7 base pairs that are repeated consecutively. Individuals may vary in the number of repeats and/or the content of the repeats. The variation in the content of the repeats occurs in either a change in the base within a repeat unit or as a deletion in the repeat unit. STRs used in forensics are either tetranucleotide or pentanucleotide repeats. STRs are highly abundant and well studied in the human genome, and their small size and the alleles' small size range facilitate typing from highly degraded, small quantities of starting material. There are 13 CODIS core loci that are being uploaded into the national DNA database.

Mitochondrial DNA. Mitochondrial DNA (mtDNA) is useful for forensic DNA in that it exists in high copy numbers in each cell, and therefore has a better chance of being detected in small samples. Moreover, mtDNA is maternally inherited, so any individual within the maternal lineage may provide an mtDNA reference sample. Finally, since the size of the amplicons (the DNA segments to be generated by PCR) is small, mtDNA can be typed from degraded DNA. Mitochondrial DNA hypervariable regions I and II are the most commonly sequenced targets in forensic DNA laboratories. The most commonly used method of mtDNA typing is fluorescent Sanger's dideoxy sequencing.

Y chromosome markers. Y chromosome markers, including Y STRs, have been recently used in many forensic DNA crime laboratories. The interest in Y chromosome markers is well supported for the following reasons:

1. The total number of male cells that are present at a rape scene may be very small in the case of rapists who are azoospermic (having no sperm) or oligospermic (having a low sperm count).

2. The total number of male cells may be low due to loss of sample or degradation.

3. Multiple semen donors may need to be identified in a multiple rape case.

4. In criminal paternity or mass disaster victim identification, determination of the haplotype of a missing individual may be conducted by typing a male relative.

5. In sexual assaults, the time-consuming and sometimes inefficient differential extraction procedure for the separation of sperm and nonsperm fractions may be bypassed.

6. Y chromosome typing may provide increased statistical discrimination in mixture or kinship analysis cases in which the discrimination obtained from autosomal markers is insufficient for identification purposes.

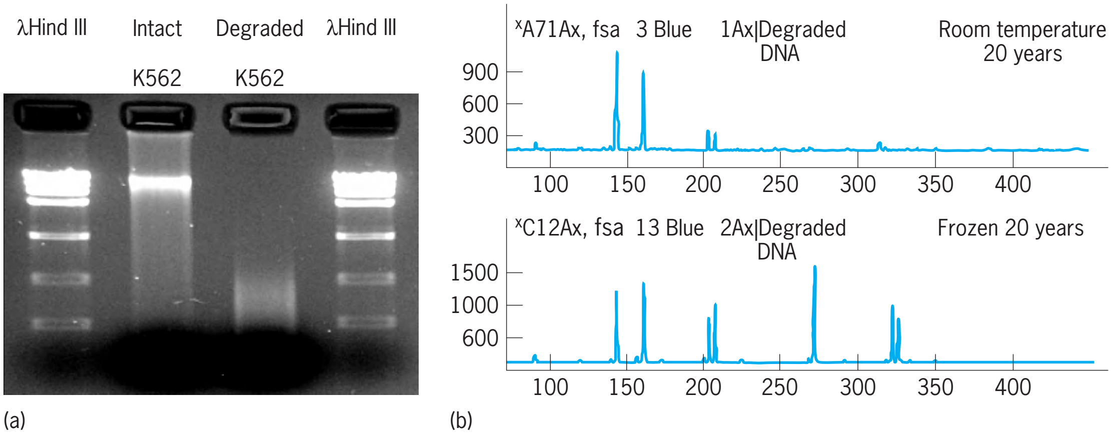
Interpretation. Once amplified and typed, the results need to be interpreted. Forensic biologists must have a clear understanding of the molecular methods utilized, with in-depth knowledge of the basis of typing (that is, the cell biology, technology, and genetics of the loci), the specific loci and amplification parameters (determined both externally and internally in developmental and internal validation studies), the empirically derived limitations of the system, the protocols and quality control measures implemented in their own laboratories, the instrument validations, the analytical software utilized in determining the DNA profiles, and the statistical and population genetics databases and software used.

In single-source samples, the interpretation requires setting a threshold of detection. That is, alleles that are at or above a certain fluorescent threshold may be designated. In highly degraded DNA samples, the interpretation requires more analysis, as higher-molecular-weight loci generally degrade more rapidly than lower-molecular-weight loci. In this case, the alleles are expected to appear in decreasing amounts as the size of the alleles increases (see **illustration**). In mixtures, the interpretation requires considering all combinations of alleles that are present, allele ratios within and among loci, and the sample type and condition.

Future technologies. Many technologies applicable to forensic biology are currently under investigation, including the following:

Hand-held, microcapillary STR typing. Development of hand-held DNA typing devices has obvious advantages for forensics in that samples might be processed quickly at the crime scene. The potential to rapidly determine the DNA type of samples left at a crime scene coupled to the growing national DNA database provides a powerful tool to law enforcement, as "cold hits" might provide leads early in the investigation. Devices have already been developed that can detect microbial pathogens and can be used out in the field. Moreover, under development are miniaturized microcapillary devices for typing human STRs.

Single nucleotide polymorphisms. Single nucleotide polymorphisms (SNPs) are single base sites that vary between individuals and as such can be used in forensic DNA typing. They have already been utilized in forensic DNA tests. The PCR-based DNA typing systems HLA-DQ Alpha and Polymarker loci were based on SNPs. The detection was by reverse dot blot



Results of degrading DNA by mechanical disruption. (*a*) Agarose gel electropherogram of intact DNA and degraded DNA, electrophoresed in lanes 2 and 3 through a 1% agarose gel, stained with ethidium bromide, and then photographed with UV light. Molecular-weight ladders of Lambda Hind III were placed on either side of the samples in lanes 1 and 4. The sample in lane 3 has been highly degraded with no detectable high-molecular-weight band as is apparent in lane 2. (*b*) Capillary electropherogram of degraded DNA stored at room temperature versus frozen. The decrease in PCR product amount as base pair size increases, as shown in the top panel, is the expected result for a degraded sample. The result demonstrates the importance of proper (frozen) sample storage.

hybridization, where the known samples were immobilized on the blots. Today there is interest in the detection of mtDNA and Y SNP typing using a variety of approaches, including primer extension assays, taq man assays, microarrays, liquid bead arrays (Luminex), and pyrosequencing.

Forensic biometrics. The ability to determine the physical characteristics of an individual by typing genetic markers has been called forensic biometrics. Inferring population of origin from DNA evidence using Y chromosome SNPs has been recently reported. The ability to determine the potential genetic origin of a perpetrator has obvious benefits to law enforcement. However, it also has obvious ethical and legal implications, as well as significant limitations, in that the genetic boundaries between races are not clear. Other applications may be in determining the age of a suspect using genetic markers.

Tissue typing using mRNA profiles. Different tissues have different genetic expression patterns. Recently the use of ribonucleic acid (RNA) has been reported for body fluid identification. The potential use of molecular technology was also reported in determining the age of a bloodstain, which could be useful in establishing the time of the crime, using analysis of messenger RNA (mRNA):ribosomal RNA (rRNA) ratios. Advantages of the mRNA-based approach, versus conventional biochemical tests, include greater specificity, simultaneous and semiautomatic analysis, rapid detection, decreased sample consumption, and compatibility with DNA extraction methodologies.

Low-copy-number amplification (LCN). Biological evidence is often found with an extremely low number of starting templates (1–15 diploid cells, <100 picograms). Analysis of LCN samples requires the development of novel strategies for its collection, analysis, and interpretation. Also known as touch sample amplification, low-copy-number amplification may require additional cycles of amplification and/or additional cleaning or sampling, but permits the analysis of extremely small amounts of sample such as might be found on the grip of a gun.

Steven Lee

Bibliography. B. Budowle et al., Forensics and mitochondrial DNA: Applications, debates, and foundations, *Annu. Rev. Genomics Hum. Genet.*, 4:119–141, 2003; B. Budowle et al., Public health: Building microbial forensics as a response to bioterrorism, *Science*, 301(5641):1852–1853, 2003; J. Butler, *Forensic DNA Typing: Biology and Technology Behind STR Markers*, 2d ed., Academic Press, San Diego, 2005; A. Carracedo and P. Sanchez-Diaz, Forensic DNA typing technologies: A review, in A. Carracedo (ed.), *Forensic DNA Typing Protocols: Methods in Molecular Biology*, pp. 1–297, Humana Press, Totowa, NJ, 2004; M. A. Jobling and P. Gill, Encoded evidence: DNA in forensic analysis, *Nat. Rev.*, 5:739–751, 2004; E. L. Jones, Jr., The identification of semen and other body fluids, in *Forensic Science Handbook, volume II*, 2d ed., pp. 329–399, Pearson Education, New Jersey, 2005; N. Rudin and K. Inman, *Forensic DNA Analysis*, 2d ed., CRC Press, Boca Raton, FL, 2001; J. A. Siegel, G. C. Knupfer, and P. J. Saukko (eds.), *Encyclopedia of Forensic Sciences*, vols. 1–3, Academic Press, 2000.

Forensic botany

The application of plant science for purposes of civil or criminal law. Forensic botany has opened aspects of the field of botany to novel avenues within criminal investigations. Today, several subdisciplines of plant science are being applied successfully in criminal investigations—for example, plant systematics including identification of plant species, plant anatomy, and plant ecology. Often, forensic applications come from combinations of these areas in a given case.

Plant systematics. Plant systematics is the study of evolutionary relationships among plants. Certain knowledge in this field is of considerable use in criminalistics and crime scene work. Being able to attach the proper scientific name to a plant is important in several situations. These items should be gathered carefully and notations made concerning exact location of the plant materials. Collection of plant material evidence is especially important in crime scene investigations, but often is overlooked or collected in an improper manner. The identification of plant materials can involve whole plants, plant fragments, and seeds. There are guides to plant identification called floras. These books are most useful when whole plants or whole flowers are to be identified. In a given setting, the more local the plant identification guide (flora), the better. For example, if dealing with a crime scene at or near Yellowstone National Park, the best source would be the Park's flora. *See* CRIMINALISTICS; FORENSIC EVIDENCE; PLANT TAXONOMY.

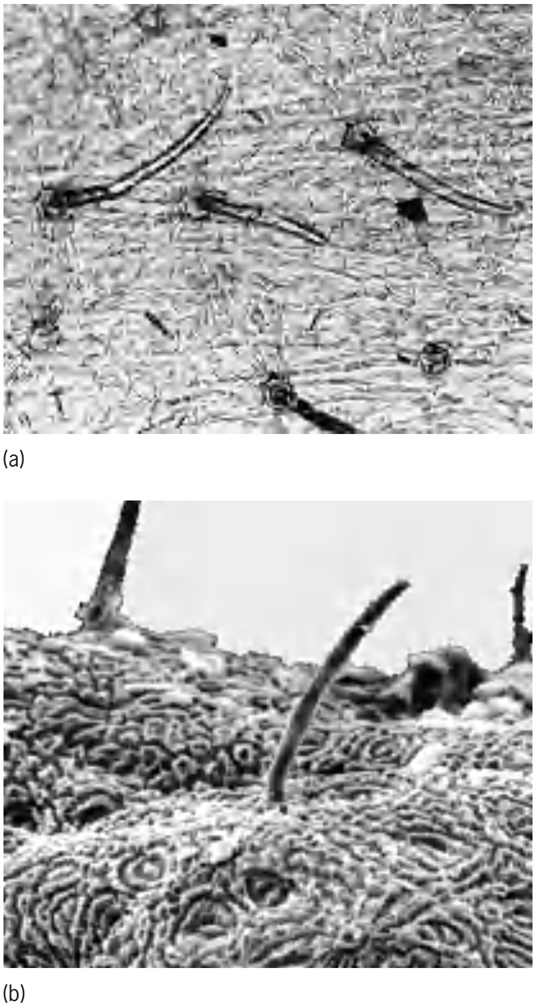
In some cases, plant fragments may be the only material available, for which there may be published guides. There are several picture guides to plant seeds. These are more difficult to use than the floras because distinctions among seeds can be very subtle. There also are guides for leaf fragments of trees and branch tips from woody shrubs and trees. With seeds and plant fragments, botanical expertise from a local college or university may be needed.

Applications of plant systematics. Sometimes homicide victims may be found in places far from the murder site. In one case, a blanket-wrapped body was found in rural grassland with assorted plants attached to the victim's clothes and the blanket. The plant fragments could not have come from that grassland because they were from ornamental plants associated with formal gardens. The person had been killed in the city and the corpse transported many miles into the country. The ornamental plants were in agreement with those found in a suspect's yard, linking the victim to the suspect, and perhaps to the crime scene.

In another case, a kidnapping victim was grabbed from the street and released somewhere in the mountains above Boulder, Colorado. The investigators suspected useful evidence might be found at the release site. A van used in the crime had been recovered. The victim identified the van, and her DNA was found in it. As investigators inspected the vehicle, they collected plant fragments and seeds from the seats and floor, the floor pads and pedals, the tire treads, the outside window wells, and the undercarriage of the van. The plant samples were collected with forceps and placed in coin envelopes. Once the plants pieces were assembled and identified, the release site was pinpointed and further evidence was found on site. Investigators are urged to collect plant material with forceps rather than with the special vacuum sweepers found in many departments, because vacuuming shatters dry plant fragments and complicates the identification process.

Plant anatomy. Plant anatomy is the study of plant cells. Applications of plant anatomy in forensic science have been based in large part on a unique characteristic of plants—their cellulose cell walls. Cellulose is a complex carbohydrate that is highly resistant to ordinary decomposition processes. For example, wood is made up of cellulose that has undergone a process known as lignification. Plant cells can pass through the human digestive tract with their walls intact, maintaining the same size and shape they possessed before ingestion. Because the human diet in any culture is made up of a relatively few kinds of plant foods, seldom more than 75 different kinds, and because each food plant has cells of distinctive sizes, shapes, cell arrangements, and cell inclusions such as crystals, it is possible to identify the plants consumed in someone's last meal (see **illustration**). This use of plant anatomy has been employed especially to postmortem the stomach contents. *See* PLANT ANATOMY.

A further use of stomach contents can be an estimation of time of death because the valve allowing stomach contents to proceed to the small intestine



Surface of a green bean pod, showing its spines using (*a*) light microscope (*courtesy of David O. Norris*) and (*b*) scanning electron microscope (*courtesy of Meredith Lane*).

closes at death. If the individual was healthy and had eaten an average-sized meal, the material would remain in the stomach for only 1–3 h under ordinary circumstances and then pass to the intestines.

Fecal analysis shows identifiable plant cells as well, a technique used with great success by anthropologists in their studies of ancient cultures.

People are sometimes stabbed with wooden instruments. One of the properties of cellulose is that it is birefringent (glows) under light microscopy. This allows determination that the stabbing was carried out with a wooden instrument, and may be used as well to link the instrument with a specific wound.

Plant anatomy applications. Sometimes the contents of the last meal can play a role in criminal investigations. In a "black widow" case, the suspect claimed to have fed the victim his last meal (supper), giving her an alibi for the estimated time of his murder. But analysis of his stomach contents showed his last meal differed greatly from her menu. He was observed earlier in the day eating his true last meal (breakfast) at a restaurant. This occurred at a time for which the "black widow" had no alibi.

In another case, the vegetables that had been selected for a pizza order helped the delivery person recall the individuals in the room where the delivery was made. Several persons in the room were killed, and the investigators thought the perpetrator might also have been present when the pizza arrived, so the deliverer's recall was significant.

Plant ecology applications. Plant ecology is the study of the relationships between plants and their physical and biological environments. The body of a missing person was discovered along a ditch in midsummer in the Great Plains wheat country. An attempt had been made to hide the corpse by covering it with giant sunflowers growing along the roadside. The victim had been reported missing only a few days before the body was discovered, but the crime scene investigators suspected the person had been dead for a longer time due to the state of decay and evidence from forensic entomology. The suspect had left the area about 10 days before the body was discovered and well before the missing person's report was filed. The sunflowers used for covering the body were wilted. In order to determine how many days it took for an uprooted sunflower to reach this particular wilted condition, several living sunflowers were collected from the site where the body was found, exposed to the environmental conditions at the site, and observed daily until the same level of wilting occurred. The appropriate level of wilting took 8 days and was maintained at that state up to 15 days, providing further evidence that the homicide had occurred before the person was reported missing and before the suspect had left the area. *See* ECOLOGY.

A federal agency suspected a kidnapping had occurred based on a single photograph it had received. This photo was circulated widely at forensic science meetings and elsewhere. This photo was all the evidence the agency had that a crime had occurred, and workers were anxious to locate the crime scene. It was concluded that the picture had been taken in the upper montane ecological zone of the southern Rocky Mountains, in either northern Colorado or southern Wyoming, based upon the Engelmann spruce trees and logs, the metamorphic rocks, and certain shrubs shown in the photo. But it was not possible to provide a more specific location.

NecroSearch International Ltd. is a nonprofit organization made up of people from many disciplines that have forensic relevance. Their purpose is to locate clandestine graves. Sequential changes in plant species known as plant succession patterns are distinct over the remains of buried bodies and often remain distinct from surrounding ground cover for many years. Evidence of plant succession patterns has been used in locating such graves. *See* ECOLOGICAL SUCCESSION.

Summary. Forensic botany is a tool in criminal investigations. Its techniques are simple for the most part and are readily accepted in court testimony. For much of the work, qualified consultants can be found at local colleges and universities among the biology faculties trained in plant science.

Jane H. Bock; David O. Norris

Bibliography. J. H. Bock and D. O. Norris, Forensic botany: An under-utilized resource, *J. Forensic Sci.*, 42(1):364–367, 1997; J. H. Bock and D. O. Norris (eds.), *Handbook of Forensic Botany*, Humana Press, 2004; M. J. Crawley (ed.), *Plant Ecology*, 3d ed., Blackwell Science, Oxford, England, 1997; W. C. Dickenson, *Integrative Plant Anatomy*, Academic Press, London, 2000; S. Jackson, *No Stone Unturned*, Kensington Publication Corp., Boston, 2002; D. O. Norris and J. H. Bock, Method for examination of fecal material from a crime scene using plant fragments, *J. Forensic Invest.*, 52(4):367–377, 2001; J. S. Sachs, *Corpse: Nature, Forensics, and the Struggle To Pinpoint Time of Death*, Perseus Publishing, Cambridge, MA, 2001.

Forensic chemistry

The application of chemistry to the study of physical materials or theoretical problems, the results of which may be entered into court as technical evidence. Forensic chemistry is part of the broader field of forensic science. The boundaries of forensic chemistry are not sharply defined, and topics that are not entirely chemical in nature are included. The scope of the analyses performed depends on the criminal, civil, and regulatory statutes of a given jurisdiction. A forensic chemist may be asked to examine debris from a fire scene for the presence of accelerant residues, to determine the species of an imported animal product for ad valorem tax classification, to demonstrate the presence of a narcotic in a pill, or to determine if a certain tool made a scratch on an illegally entered door.

A forensic laboratory may be part of a governmental agency or may be within private industry. The extent of its service varies considerably. Full-service laboratories analyze a wide range of evidentiary items.

Regional laboratories usually screen cases, performing basic examinations, while the more complex tests are done in a central laboratory. Specialty laboratories, such as a medical examiner's lab or a regulatory lab, have a narrower range of services covering a few specified areas. The crime laboratory is the best known of the three types.

Evidence types. Every material in existence could become evidence at some time. Some of the items most often encountered in crime laboratories, and the information sought, are (1) body fluids and viscera to be analyzed for poisons, drugs, or alcohol, the quantitation of which may assist in determining the dosage taken or the person's behavior prior to death; (2) licit and illicit pills, vegetable matter, and pipe residues for the presence of controlled substances; (3) blood, saliva, and seminal stains, usually in dried form, to be checked for species, type, and genetic data (DNA); (4) hairs, to establish animal or human origin (and if human, the race, body area of origin, and general characteristics); (5) fibers, to determine type (animal, vegetable, mineral, or synthetic), composition, dyes used, and processing marks; (6) liquor, for alcoholic proof, trace alcohols, sugars, colorants, and other signs of adulteration; (7) paint, glass, plastics, and metals, usually in millimeter-sized chips, to classify and compare to known materials; (8) inks on documents, to determine type, dye content, or possible age; also chemical obliterations and restoration of charred documents; (9) damaged floppy disks or hard drives, for recovery of erased computer files; (10) debris from a fire or explosion scene, for the remains of the accelerant or explosive used; (11) materials seized from a clandestine laboratory, to determine what illicit drugs were being made; (12) broken headlight filament, to determine whether it was "on" or "off" at the time of an accident.

Regulatory labs receive any product that has defined legal specifications. These may include chemical composition, physical performance, or the absence of hazardous chemicals. Private laboratories act as consultants for either side in a legal contest by providing general and specific chemical knowledge.

Legal aspects. Forensic chemists are concerned with more than just the technical aspects of an analysis. They are also involved with the handling of evidence before and after it enters the laboratory. An item considered to be evidence goes through five stages: recognition, collection, preservation, analysis, and testimony. At each stage there are legal requirements that must be met if the item is to be judged admissible in court, and chemical precautions which must be observed if it is to be analyzed properly. One of the legal requirements is that the chain of evidence, that is, the origin, transfer, possession, and analysis of each item, must be recorded as it occurs. All packages and seals should remain intact. Detailed notes are necessary.

At some crime scenes the chemist, with the advantage of specialized knowledge, is aware of the potential use of certain materials involved and can recognize, recover, and preserve evidence so that its chemical usefulness is not ruined. Communication and cooperation between field assistants and laboratory analysts allow the problems to be solved rapidly and ensure that the evidence is received in good condition.

For court testimony, a forensic chemist must be able to explain complex instrumental techniques to a nonspecialist jury in an understandable manner, and must have a thorough knowledge of each instrument and chemical property in order to withstand cross-examination.

Purposes of analyses. Before beginning an analysis, the forensic chemist must understand what the investigator or prosecutor or defense is seeking from it. To prove possession of a controlled substance, the chemist must identify a compound according to the definition in the relevant statute. If a drug statute specifies the difference between possession of cocaine base (crack) and cocaine as a salt (coke), then the analytical procedure must be able to identify each species separately. In general this is referred to as establishing an element of a crime. Another purpose that rules the analytical process is proving or disproving a statement made by a suspect, victim, or witness. Most of these analyses are of the "either-or" type. If a red car is struck during a hit-and-run incident and the police locate a car with red smears, the question is, "Could the red smears on the suspect's vehicle have come from the victim's car?" If the suspect should offer an alibi that the red paint came from the trim of his garage door, then a second question arises, "Can the laboratory show this statement to be true or false?"

Databases. Sometimes the laboratory is asked for investigative help: "What can the physical evidence tell an investigator that will help him find a suspect?" The answer may be a simple description of the evidence such as that the suspect vehicle has a metallic green top coat and shows damage to the right front headlight and grille area. In other cases, searchable databases are consulted to identify individuals or to link crimes.

In the United States, the National DNA Index System (NDIS) contains DNA profiles from convicted offenders. All states collect DNA samples from convicted sex offenders; some states have broadened their collections to include all convicted felons. The Federal Bureau of Investigation (FBI) manages the national index, which is a combination of all the state data. DNA from blood, semen, saliva, skin cells, and so on can be profiled and compared to the national index to identify a suspect. In addition, serial crimes can be linked by a common profile. The agencies involved can share their case information to help identify a suspect.

The Integrated Automated Fingerprint Identification System (IAFIS) allows laboratories to search a single latent print against state and national (FBI) collections of 10-print cards obtained during arrest procedures. *See* FINGERPRINT.

Drugfire and IBIS (Integrated Ballistics Identification System) are systems that allow for searching of breech markings on cartridge casings and striations on bullets. They can link shooting incidents even without access to the weapon itself.

Theory of analysis. Forensic examinations are designed for one of two purposes: identification or comparison. In the former case, a particular chemical species, such as blood alcohol level, the amount of a drug in a tissue, or the amount of barium and antimony in gunshot residues, is identified and often quantitated. DNA analyses by short tandem repeats (STR) have now reached a level of discrimination such that people can be uniquely identified (except identical twins).

Comparative examinations determine consistencies or inconsistencies between the chemical and physical properties of two samples in an attempt to show or exclude a common origin. Many of the materials encountered in a forensic laboratory are manufactured in large batches that have a homogeneous nature. A thousand gallons of paint may be used on many cars. A sample from any of these cars will be consistent with any other sample. If a car is involved in an accident, the paint can be matched to it, but not to the exclusion of the other cars. Inconsistencies are positive, since they exclude the possibility of a common origin.

Thus, for each item of evidence, an analytical scheme can be devised which leads to identification or comparison. An example of identification involves drug analysis, where it is necessary to identify the controlled substance and quantitate it. A small portion of a pill is submitted to a series of color tests. The results indicate the class of drug presence (barbiturate, alkaloid, and so on), but not which member of the class. The drug is then extracted from the excipients in the pill, and individual members of the class are separated by a chromatographic method. Confirmation of the identity is typically obtained from infrared spectroscopy and quantitation by gas chromatography. For purposes of comparison, the paint from a suspect car is compared to paint from a hit-and-run victim. The color, and number and relative thickness of the layers are studied; the chemical behavior toward solvents and spot-test reagents is observed; the inorganic pigments are identified and the polymer film characterized. If both paints have similar results in all tests, they are consistent; but the report must state that another source could exist.

In the analytical scheme, modification may be made as needed. When the sample is small, nondestructive tests are done first. These tests include microscopic examination, index of refraction, density, x-ray diffraction, x-ray fluorescent elemental analysis, and in some cases, infrared spectroscopy.

Equipment and instrumentation. A modern forensic laboratory has a vast array of sophisticated instruments. Each has its own purpose, and is used for several different evidence types. One of the simplest tests is the spot test, which gives a color change, a precipitate, gas evolution, or a crystal when it reacts with a particular class of chemicals. Spot tests are not unique, but allow for the rapid screening of many samples. Other simple tests are solubility in water and organic solvents, density, index of refraction, melting and boiling point, and flash point.

Microscopy. Probably the most frequently used piece of equipment is the microscope. Low-power stereomicroscopes are used to examine the surfaces of opaque materials such as paints, soils, fibers, and paper. Examples of the uses of optical transmission microscopes include the examination of the internal structure of hairs, measurement of the index of refraction of small chips of glass, checking for spermatozoa in assault cases, and the study of tissue samples in pathology laboratories. Optical transmission microscopes may also have accessories for luminescent studies, dispersion staining, and the measurement of optical properties with polarized light. Comparison microscopes allow the chemist to view two samples side by side on a split screen. Metallographic microscopes are used to study fractured metal parts in liability investigations. *See* MICROSCOPE; OPTICAL MICROSCOPE.

The scanning electron microscope (SEM) with an elemental analyzer also has found forensic applications. Magnifications up to 100,000 times can be achieved with a greater depth of field than in optical microscopes. Minute marks on fired casings, scratches on tumblers, and the composition of very small paint layers have been compared using the scanning electron microscope. Research has shown that minute particles (less than 0.001 mm) of a certain shape and elemental composition are unique to gunshot residue, and the scanning electron microscope can be used to examine material from the hands of individuals for such residues. *See* SCANNING ELECTRON MICROSCOPE.

Spectroscopy. A second common group of instruments are in the field of spectroscopy. The infrared spectrum of a pure compound is a chemical fingerprint. This method is used to positively identify a substance, and can also classify materials present in mixtures. Fourier transform infrared (FTIR) instruments have supplanted dispersion instruments. Ultraviolet-visible spectrophotometry is a screening technique that characterizes materials that possess color.

Elemental analyses are done by using either a graphite furnace atomic absorption spectrophotometer (AAS) or ion coupled plasma (ICP) emission spectrograph, each capable of determining the major or trace elements in a sample, with ultimate detection levels in the parts per billion being common. Applications include the identification of metallic poisons such as lead and arsenic, trace pollutants, and gunshot residue.

The mass spectrometer, used in drug analysis since it can positively identify a volatile chemical substance, is often coupled to a gas chromatograph as a specific detector. Toxicology labs rely upon these instruments for rapid screening at low levels of occurrence. Special methods have been developed for the identification of explosives. Trace chemists use mass spectrometry to identify petroleum products used as accelerants and as a detector on an ICP emission spectrometer. *See* FORENSIC TOXICOLOGY; MASS SPECTROMETRY; SPECTROSCOPY; TOXICOLOGY.

Chromatography. Since most evidence is a complex mixture containing both relevant and irrelevant materials, forensic chemists depend upon separation techniques. The classical methods of extraction, distillation, and precipitation are sometimes used, but they have been replaced to a large degree by chromatographic methods. Chromatography not only separates but also quantitates at the same time.

The use of specific detectors has broadened the use of these methods. Thin-layer chromatography is a rapid screening technique for drugs, explosives, poisons, and ink dyes, while gas chromatography is both a screening and quantitating technique. These methods are applied to drugs, arson accelerants, pesticides, and hazardous chemicals. Liquid chromatography, which is a relatively new forensic tool, has the potential to replace many thin-layer and gas chromatography procedures, since fewer preparations of the sample are required, more versatile detectors can be used, and nonvolatile samples can be examined. Drugs, inks, explosives, and carcinogenic materials have been detected by liquid chromatography. *See* CHROMATOGRAPHY.

Electrophoresis. Electrophoresis separates large molecules such as proteins and enzymes. Forensic serologists can study these species in dried blood stains and compare them to standard samples from a suspect. Most crime laboratories have stopped performing conventional blood group analyses and have switched to DNA analyses. One procedure uses capillary electrophoresis to separate alleles. Since the frequency with which each allele occurs in a population is known, the probability of two persons having the same DNA profile can be calculated. *See* ELECTROPHORESIS.

X-ray techniques. These techniques are widely used in forensic laboratories. X-ray diffraction can positively identify any crystalline material, even in mixtures. Paint pigments can be run without separation from the film, since the film is noncrystalline and does not diffract the x-rays. Other uses are for soils, minerals, household chemicals, and drugs. X-ray fluorescence techniques can be used to determine the amounts of elements heavier than sodium, simultaneously, in 100 seconds. Such a rapid method has found acceptance in many laboratories. *See* X-RAY DIFFRACTION; X-RAY FLUORESCENCE ANALYSIS.

In addition to the above instrumental techniques, many less common physical and chemical examinations are conducted by forensic laboratories. It is incumbent upon the analyst to choose those which will satisfy the legal requirements for a specific case. *See* ANALYTICAL CHEMISTRY; CLINICAL PATHOLOGY; FORENSIC MEDICINE.

Michael J. Camp

Forensic engineering

Engineering applied toward the purposes of law. In forensic engineering, the majority of investigations are carried out in the context of civil litigation; for example, the cause of a plane crash is often due to defective design, which is a civil litigation consideration. However, the crash could be due to a bomb or terrorist activity in the cockpit, which is a criminal consideration. Often the cause of an accident is initially unknown or misattributed. *See* CRIMINALISTICS.

Accident reconstruction. Accident reconstruction is the a posteriori scientific process of using the facts of an accident and the appropriate natural laws of science to determine the possible circumstances and causes, consistent with the available data. The quality and reliability of the reconstruction depend on the amount and quality of the available data, as well as the skill and knowledge of the investigator.

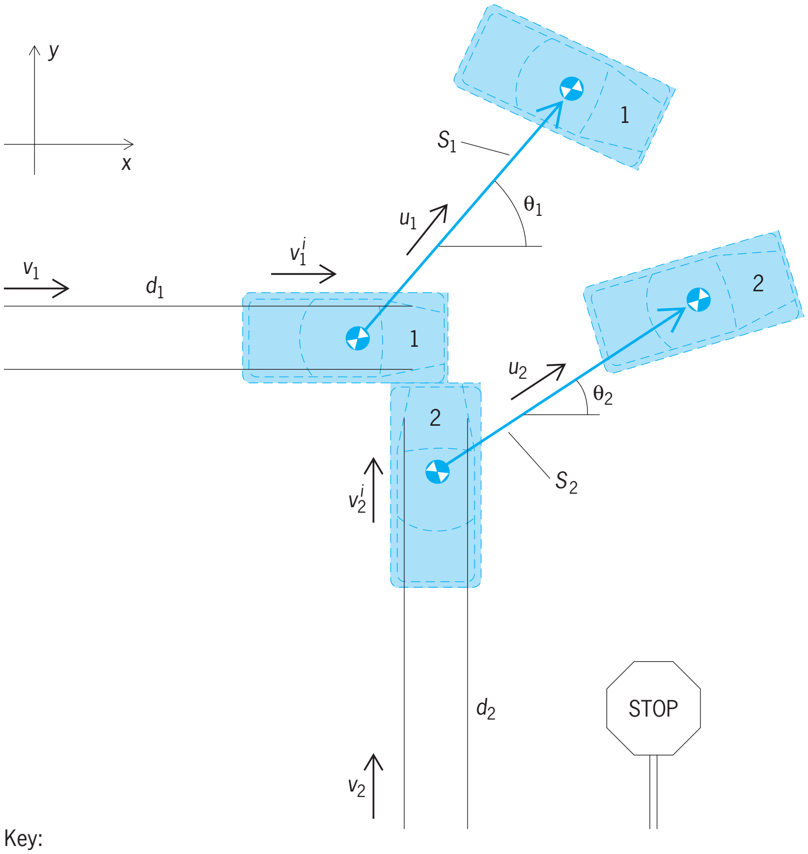
Consider, for example, a two-vehicle, right-angle, intersection collision (see **illustration**). The two vehicles had proceeded through an intersection, and then an impact occurred. Each vehicle had skidded prior to impact, and the final positions of the vehicles are indicated relative to the point of impact. For an analysis, the input data required (distances and angles as shown along with vehicle weights) would have to be obtained from a field investigation and from databases of vehicle properties. The principle of conservation of linear momentum, which follows from a straightforward application of Newton's laws, where all calculation details are given, enables the engineer to calculate the speeds of both vehicles at impact (v_1^i, v_2^i) and at the onset of preimpact skidding (v_1, v_2), based solely on the data (independent of driver or witness statements). Depending on the actual input data and the calculated speeds, the results may be used in a variety of ways. Driver 2, for example, may claim that he stopped at the stop sign, and the accident occurred because driver 1 was speeding through the intersection. Based on the input data for the case illustrated, this claim cannot be true: the accident occurred because driver 2 did not stop at the stop sign, and driver 1 was not speeding. *See* NEWTON'S LAWS OF MOTION.

Product liability. In a product liability investigation, engineers are asked to determine whether a product is defective and whether the defect was causally related to any injuries that may have occurred. This is a challenging area of investigation and analysis. The three types of product defects recognized by law are (1) design defects (the product lacks those elements necessary for its safe and foreseeable uses), (2) manufacturing defects (the product was not made according to the manufacturer's specifications), and (3) failure to warn or instruct (the product was not accompanied by adequate warnings or instructions for its proper and safe use). In a lawsuit, if a forensic engineer has been retained by the plaintiff (the person bringing the lawsuit), the engineer's role is to objectively analyze the product for defects, which may or may not be found, and to causally relate the plaintiff's injuries to the product defect. If a forensic engineer has been retained by the defendant, the engineer's role will be to objectively defend the product, if it can be done, in view of the defect allegations made by the plaintiff. As part of the defense, the engineer may opine that the alleged defect does not exist or that the injuries sustained by the plaintiff are not causally related to the product design. Product liability investigations can be very intense and can involve considerable analysis and calculations, testing, and computer modeling, often requiring the expenditure of large sums of money by both the plaintiff and defendant.

Design defects. As a simple but meaningful example of a design defect, consider cigarette lighters. Children are naturally inquisitive and may pick up lighters and try to light the flame. Some lighters have bright colors and shapes that make them particularly attractive to children. If the lighter does not have a child-proof or -resistant mechanism, which prevents children from lighting the flame, the lighter is defectively designed. The plaintiff's engineer in this case would show that at the time of manufacture it was practical and feasible to have provided a child-proof mechanism, and that the device may have even been in existence on other lighters, which would have prevented the accident.

Manufacturing defects. The fact that a person is injured using a product does not necessarily mean the product is defective. In addition, defective products do not cause injuries every time they are used. For example, a knife is designed to have a sharp edge in order to cut. If the user is cut, it does not necessarily follow that the knife is defective. However, suppose that a person is using a knife to carve a turkey and the blade fractures, causing the person to lose balance, to fall forward, and to be stabbed by the fractured blade. After the accident, a metallurgical analysis of the knife blade reveals that the steel was not properly heat-treated and that there were inclusions in the blade, which weakened it and caused the failure; this determination indicates that the knife was defectively manufactured.

Failure to warn. Car air bags can be life-saving devices if properly designed and used. When an air bag deploys in an accident, its front can move at up to 200 mi/h (320 km/h). A person sitting too close to the air bag can be seriously injured or killed by the impact to the head or upper torso. Small adults and children are particularly susceptible, as well as infants in rearward-facing child seats. Automobiles carry prominent warnings that occupants should not sit too close to air bags, that small children should sit in the rear, and that rearward-facing child seats should never be placed in the front seat of an air bag–equipped vehicle. Without such a warning, the vehicle would be defective by virtue of a failure to warn. The plaintiff's biomechanical engineer in this case would show by tests and/or calculations, or even using the automobile manufacturer's own data, that a person or child sitting too close to the deploying air bag would be subjected to injurious levels of impact forces.



Key:

d = preimpact skidding distance
S = postimpact movement
v = velocity at beginning of skid
v^i = impact velocities
u = postimpact velocities
θ = angles as shown

Intersection collision.

Biomechanics of injury and death. Engineering science and methodology is being applied to injury and death investigation, a subject rooted in human biology. Biomechanical engineers working in the field of injury and death investigation are generally interested in applying engineering principles to the understanding of injury causation mechanisms in accidents, and in developing the criteria and conditions for when injury is likely to occur. For an accident, the kinematics of movement of the human body must be calculated and then correlated with any impacts and injuries that the body may have sustained. In an automobile accident, the movement of an occupant with respect to the vehicle (occupant kinematics) is often correlated with the acceleration-time history of the occupant, the occupant's change in velocity (delta-v), and the resulting injuries. This is a critical step if the occupant alleges an injury because of the accident or a vehicle defect. A very active research field in biomechanics is the determination of criteria that delineate the conditions in which injury may occur in a particular accident. *See* BIOMECHANICS.

Computer forensics. This rapidly changing area of forensic engineering refers to methods for solving crimes committed using computer technology. Such crimes include extortion, violations of Internet security, hacking into supposedly secure Web sites and computers, computer theft of sensitive information and proprietary files (including national security data and financial data), and creation and propagation of computer viruses. Computer forensics uses techniques and procedures for recovering magnetic data, tracking hackers, and analyzing audio tapes, videotapes, and photographs for possible tampering and alteration. *See* DIGITAL EVIDENCE; INTERNET.

Cause and origin. This area refers to investigations to determine the cause and origin of fires and explosions, and to determine if the cause was accidental (electrical shorting, product defect, cooking mishap, and so on) or intentional (arson or planting an explosive device). An understanding of heat transfer, burn patterns, combustion rates, pressure-wave propagation, and analysis of explosive residues is essential for cause and origin investigators.

Structural collapse, blast loading. This area refers to the catastrophic failure of structures either during construction or after the structure is in service. For example, the Hyatt Regency Skywalk collapse in Kansas City, Missouri, on July 17, 1981, was caused by a simple design error and resulted in the loss of 114 lives. The collapse was caused by an improperly designed connection, which did not meet Building Codes, where the vertical hanging rods passed through the box beams that supported the pedestrian walkways. Forensic civil engineers are also concerned with understanding the effect of blast and seismic loading on structures in order to ensure that future structural designs are safer and more resistant to dynamic loading arising as a result of explosions and earthquakes. *See* EARTHQUAKE ENGINEERING.

Significant structural lessons were learned from the bombing of the Murrah Federal Building in Oklahoma City on July 19, 1995, and its resulting collapse; these lessons (outlined in the Proceedings of the 1st Congress of the American Society of Civil Engineers in 1997) will be used to mitigate blast loading effects on future reinforced concrete building designs. Forensic engineers are becoming more actively involved in developing counterterrorism tactics and procedures, and this activity will likely grow as a separate area of forensic engineering. *See* STRUCTURAL DESIGN.

Steven C. Batterman; Scott D. Batterman

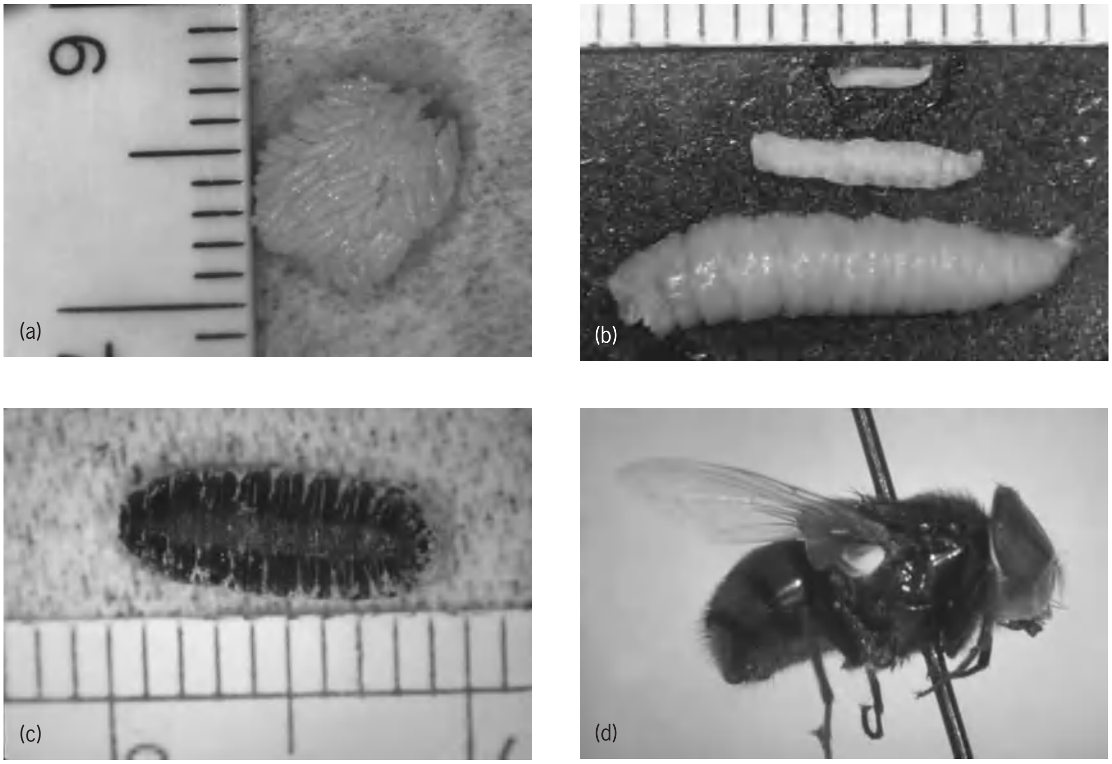
Bibliography. S. C. Batterman, *Education of Forensic Engineers*, 20(6), American Academy of Forensic Sciences, 1990; S. C. Batterman and S. D. Batterman, Accident Investigation/Motor Vehicle (accident reconstruction and biomechanics of injuries), in J. A. Siegel et al. (eds.), *Encyclopedia of Forensic Sciences*, Academic Press, 2000; S. C. Batterman and S. D. Batterman, Forensic engineering, chap. 12 in Y. H. Caplan and R. S. Frank (eds.), *Medicolegal Death Investigation: Treatises in the Forensic Sciences*, 2d ed., Forensic Science Foundation Press, 1999; F. P. Beer and E. R. Johnston, *Vector Mechanics for Engineers: Statics and Dynamics*, McGraw-Hill, 1977; D. T. Greenwood, *Classical Dynamics*, McGraw-Hill, 1977; K. L. Rens (ed.), *Forensic Engineering: Proceedings of the 1st Congress*, American Society of Civil Engineers, 1997; H. Yeh and J. I. Abrams, *Principles of Mechanics of Solids and Fluids: Particle and Rigid Body Mechanics*, McGraw-Hill, 1960.

Forensic entomology

The application of insect evidences or damage for purposes of civil or criminal law. Forensic entomology includes three subdisciplines: stored product entomology, structural entomology, and medicolegal entomology. Stored product entomology deals with damage to stored products, such as food materials or clothing, caused by insect activities. Structural entomology deals with damage to buildings and other structures by pest species such as termites or carpenter ants. Both these areas typically involve civil actions when there is a need to assess monetary damages resulting from insect activity. Medicolegal entomology deals with insects as evidence in criminal events, most frequently homicides. This is the area most commonly recognized by entomologists, law enforcement, and the public as forensic entomology. The application of entomological evidence to criminal investigations is not new, with the first recorded use coming from twelfth-century China. Beginning in the early 1980s, there has been a resurgence of interest in the field, and new applications of insect evidence are being exploited. *See* CRIMINALISTICS.

Insect evidence. Insects can serve as evidence in criminal investigations in several different ways. They can provide data for estimating the time since death (postmortem interval), detecting possible movement of the body following death and circumstances of the crime scene, and assessing antemortem versus postmortem wounds on the body. In addition, insect evidence can provide alternative specimens for toxicological analyses in highly decomposed bodies, yield DNA material linking a suspect to a victim for ectoparasitic taxa, and help assess periods of abuse or neglect of children and the elderly. Of these, the most frequent application is the estimation of the postmortem interval. *See* DEATH.

A dead body provides a changing and ephemeral food source for a wide variety of organisms, ranging from bacteria to vertebrate scavengers. Among these organisms, insects are major factors because they arrive at decomposing remains in a predictable pattern and develop in known patterns. Insects arrive shortly following death, often within minutes. The first arrivals are most frequently flies in the families Calliphoridae (blowflies and bottle flies) and Sarcophagidae (flesh flies). The female flies arrive and begin to investigate the natural body openings of the head, anus, and genitals. Wounds present on the body provide another area for activity. The female flies take a meal of blood or other exudates from the body and then deposit their eggs or, in the case of the Sarcophagidae, larvae into these body openings. This action starts a biological clock which is stopped and interpreted by the forensic entomologist when the body is discovered and the insects are collected. The use of insects to estimate the postmortem interval requires an understanding of the insect's life cycle, the relationship of the insect to the remains, and the relationship of the remains to the habitat in which they are discovered.



Chrysomya megacephala **(family Calliphoridae). (*a*) Egg mass. (*b*) Larval instars: first (top), second (middle), and third (bottom). (*c*) Puparium. (*d*) Adult female.**

Insects pass through a number of distinct life-cycle stages. Using the blowfly in the family Calliphoridae as an example, the female fly will arrive at the body and deposit eggs (**illus.** *a*) in the natural body openings or in wounds. These eggs will hatch into the larvae (maggots) which will feed on the decomposing tissues. There are three distinct larval stages (illus. *b*), called instars, with a molt between each stage. Once the maggot has reached complete development, it will cease feeding and move away from the remains to enter the pupal stage (illus. *c*), an inactive stage during which the larval tissues are reorganized to produce the adult fly (illus. *d*).

The insects encountered on a corpse in any given habitat will consist of some species unique to that particular habitat and some species having a wider distribution. The former group may be restricted to a particular geographic area or a particular habitat type within a given geographic area. For example, in Hawaii there are taxa that are restricted to a rainforest habitat, while others are specific to a more arid tropical habitat. Taxa having wider distributions are frequently encountered in several different habitat types and are typically highly mobile species, such as those taxa that are found both in rainforests and on arid beaches. Many of those taxa closely tied to carrion show this wider pattern of distribution. In estimating the postmortem interval, taxa from both groups may, under given circumstances, provide essential information concerning the history of the corpse. *See* INSECTA.

Species analysis. Of those insect taxa having a direct relationship to the corpse, there are four generally recognized relationships:

1. Necrophagous species actually feed on the corpse. Included are many of the Diptera (Calliphoridae and Sarcophagidae) and Coleoptera (families Silphidae and Dermestidae). Species in this group may be the most significant isolatable taxa for use in postmortem interval estimates during the earlier stages of decomposition, defined here as days 1–14. *See* DIPTERA.

2. Parasites and predators of necrophagous species constitute the second most significant group of carrion-frequenting taxa. Many of the Coleoptera (Silphidae, Staphylinidae, and Histeridae), Diptera (Calliphoridae and Stratiomyidae), and Hymenoptera parasites of Diptera larvae and puparia are included. In some instances, Diptera larvae, which are necrophages during the early portions of their development, become predators during their later larval development. *See* COLEOPTERA.

3. Omnivorous species include taxa such as ants, wasps, and some beetles, which feed on both the corpse and associated arthropods. Large populations of these may severely retard the rate of carcass removal by depleting populations of necrophagous species. *See* ANT; ARTHROPODA; WASP.

4. Adventive species include those taxa which use the corpse as an extension of their own natural habitat, as in the case of the Collembola, spiders, and centipedes. Acari in the families Acaridae, Lardoglyphidae, and Winterschmidtiidae, which feed on molds and fungi growing on the corpse, may be included in this category. Of less certain association are the various Gamasida and Actinedida, including the

Macrochelidae, Parasitidae, Parholaspidae, Cheyletidae, and Raphignathidae, which feed on other acarine groups and nematodes. In these investigations, what is being estimated is actually a period of insect activity rather than an actual time since death.

Interpretation. During the early stages of decomposition, the estimate of the postmortem interval is most frequently based on the development rates of individual species of Diptera, most frequently the flies in the family Calliphoridae. The most mature specimens collected from the body are preserved and identified. Given the particular species and stage of development represented, the entomologist can determine the period of time required to reach that stage. As the insects are dependent on environmental conditions, particularly temperature, for their rate of development, this must be factored into the estimate. Data obtained from detailed laboratory studies of life cycles conducted under controlled conditions can be correlated with conditions at a crime scene using weather data from National Oceanic and Atmospheric Administration (NOAA) stations and the concept of accumulated degree hours or accumulated degree days, a technique originally developed to predict pest outbreaks in agriculture. This calculated time period will represent the minimum postmortem interval. It is important to note that the time period is the period of insect activity and not the actual postmortem interval. Under most conditions, these will be quite similar, but there may be factors that delay the onset of insect activity, such as wrapping or concealment of the body.

After the first 2–3 weeks, depending on the location and environmental conditions, those flies in the initial wave of invaders will have completed their development and departed the body. At this point, the estimation of the postmortem interval is based on the ecological succession of insects onto the body. Those insects initially colonizing the body change the body by their activities. These changes make the body attractive to another group of insects, which feed and change the nature of the body, thus making it attractive to yet another group of insects. This process continues until the resource of the body has been completely exhausted. By making complete collections of the insects and other organisms present on a decomposing body and comparing these taxa with results of detailed decomposition studies conducted in similar habitats and areas, periods of time during which the death most probably occurred can be determined. Care must be taken in applications of results from decomposition studies as even geographically proximate localities may support different arthropod populations. These differences in species composition and relative abundance may serve to alter the successional picture. Generally speaking, the closer to the time of death, the more accurate the estimated postmortem interval will be. The estimate begins in terms of hours and proceeds to days, months, and finally seasons of the year.

Entomological evidence is not present in all cases and, when present, may not ultimately prove to be of major significance. When insect evidence is present and properly interpreted, it may prove to be a powerful tool for the solution of the case. M. Lee Goff

Bibliography. J. H. Byrd and J. L. Castner (eds.), *Forensic Entomology: The Utility of Arthropods in Legal Investigations*, CRC Press, Boca Raton, 2000; E. P. Catts and N. H. Haskell (eds.), *Entomology and Death: A Procedural Manual*, Joyces Print Shop, Clemson, SC, 1990; M. L. Goff, Estimation of postmortem interval using arthropod development and successional patterns, *Forensic Sci. Rev.*, 5:81–94, 1993; F. Introna and C. P. Campobasso, *Entomolgia Forense: II ruolo dei ditteri nelle indagini medico legali*, Essebiemme Editore, Italy, 1998.

Forensic evidence

Scientific, technical, or other specialized knowledge applied in courts of justice. Courts are finding that some areas of knowledge that have been considered forensic sciences for decades are not based completely on scientific knowledge under contemporary legal and academic standards, but rather are based on technical or other specialized knowledge that must be evaluated under standards not clearly defined. Indeed, some of the most interesting issues facing forensic evidence arise in cases where the defining characteristics of “scientific,” “technical,” and “other” specialized knowledge are at issue.

Federal law in the United States now requires trial courts to assure that forensic evidence is reliable and valid—and not merely generally accepted by others in the witness’s field—before it is admitted in court. The highest courts of several states have also adopted this federal standard.

Historical context. Several areas of forensic specialization matured in the United States under the “general acceptance” standard announced in *Frye v. United States*, 293 F 1013 (D.C. Cir., 1923). This standard allowed scientific evidence to be admitted in court if the evidence was generally accepted as reliable by others in the expert’s field. In 1993 in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, the U.S. Supreme Court rejected the *Frye* standard as the sole test for admissibility in federal courts, and interpreted the Federal Rules of Evidence to require proof that scientific evidence was valid before it would be found sufficiently reliable to be admitted into evidence. The Court noted that “scientists typically distinguish between ‘validity’ (does the principle support what it purports to show?) and ‘reliability’ (does application of the principle support consistent results?),” but stated that “our reference here is to *evidentiary* reliability—that is, trustworthiness. . . . In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*.” In 1999, in *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, the Supreme Court made it clear that *Daubert*’s requirements of validity and reliability apply not only to scientific evidence but to all types of forensic evidence, and stated that forensic evidence must have “a reliable basis in the knowledge and experience of the witness's discipline."

Many types of forensic evidence used in criminal proceedings were developed to meet the practical needs of law enforcement and prosecutorial functions of government, and not specifically to address theoretical or empirical questions. Courts are finding that many areas of forensic specialization are not grounded in science or scientific method, and are just beginning to formulate questions that may allow empirical investigation of the validity and reliability of their fields of knowledge. Areas of forensic knowledge at this stage of development are easy targets for challenges to reliability and validity in litigation.

Areas of forensic specialization that are grounded in science and scientific method are open to challenge on any aspect of theory or practice that may affect the reliability or validity of the evidence. Testimony of a witness or a peer that an established theory or technique has been applied in a manner that is generally accepted in the field is no longer sufficient to allow the evidence to be admitted. The trial court may consider general acceptance in evaluating the reliability and validity of the evidence, but independent proof of reliability and validity is also required.

Validity and reliability. The shift from "general acceptance" to "validity" and "reliability" is guided primarily by three elements of Federal Rule of Evidence 702 and state counterparts. This rule requires that (1) admissible forensic evidence must consist of "scientific, technical, or other specialized knowledge" (2) that will "assist the trier of fact to understand the evidence or to determine a fact in issue," and (3) must be provided by a witness who is qualified as an expert by "knowledge, skill, experience, training or education."

Specialized knowledge. *Daubert* defined the "specialized knowledge" element to require "scientific validity" as a basis for "evidentiary reliability." Applied to nonscientific evidence, *Kumho Tire* requires forensic evidence to have "a reliable basis in the knowledge and experience of the witness's discipline." Both *Daubert* and *Kumho Tire* identified four factors that trial courts may use to evaluate the validity and reliability of forensic evidence: (1) whether the knowledge or technique can be (and has been) tested, (2) whether it has been subjected to peer review and publication, (3) the technique's known or potential error rate, and (4) whether it is generally accepted in the witness's community. The Supreme Court cautioned in *Daubert* that this was not a "checklist" and that trial courts could consider anything that addressed the validity and reliability of the evidence. Later, in *Kumho Tire*, the Supreme Court stated that the trial judge has discretion to decide which, if any, of the four factors to employ in evaluating the reliability and validity of forensic evidence. The range of possible challenges to reliability and validity under these standards is limited only by the creativity of litigants and proponents of areas of specialized knowledge in our increasingly complex society.

Intradisciplinary knowledge. A forensic expert's knowledge of the scope and limits of the area of specialized knowledge claimed by the witness is, of course, subject to cross-examination under any or all of the four *Daubert* factors, and other factors raised by counsel and accepted by the court. A witness should therefore be knowledgeable of authoritative texts, treatises, standards, and training materials that constitute the knowledge base of the witness's claimed area of expertise. Opposing counsel may produce these materials and ask the witness to identify those deemed authoritative in the witness's field, and then request the witness to explain any discrepancies between the witness's testimony or practices and the content of the authoritative treatises. The witness may also be asked to provide a rational, logically and substantively valid explanation of the theories, techniques, standards, customs, and practices of the witness's field of specialization, and how these were applied to the facts of the case to reach the witness's opinions or conclusions. The purpose of this type of cross-examination is to identify any acts, omissions, biases, prejudices, or other matters that may affect the validity and reliability of the witness's evidence, and to provide a basis for counsel to assure that the evidence offered by the witness is within the scope of the witness's expertise.

Extradisciplinary knowledge. More importantly, a witness should attempt to be prepared for cross-examination concerning areas of knowledge that may be beyond the witness's area of specialization but are relevant to the determination of the validity and reliability of the witness's opinions or conclusions. *Daubert*'s expansion of the judicial inquiry beyond the realm of general acceptance in the witness's community supports this type of challenge, and opposing counsel may rely on this attribute of *Daubert* to test the limits of a witness's knowledge.

For example, the definition of "science" is at issue in an increasing number of cases. A witness from a field of specialized knowledge that is claimed to be a science may be challenged not only concerning the validity and reliability of evidence as viewed from within the witness's field of knowledge, but also concerning the validity and reliability of the witness's evidence as viewed from the field of science. Cross-examination may, for example, address how the witness's evidence accords with definitions of "science" and "scientific method," or with definitions of particular areas of forensic science or forensic specialization. *See* SCIENCE; SCIENTIFIC METHODS.

Criteria for "science" may be found in the vast literature in this field. For example, nine guideposts that courts and attorneys may use to prepare and challenge scientific evidence have been proposed [B. Black and P. Lee, 1997]: (1) the explanatory power of the evidence to explain, clarify, and predict relationships in general and in the case at hand; (2) logical consistency of the evidence; (3) falsifiability; (4) the variety and severity of testing of hypotheses on which the evidence is based; (5) precision and objectivity of the reasoning and conclusions, rather than testimony consisting of broad generalizations and subjective impressions; (6) post-hypothesis testing, that is, whether the testimony is based on an ad hoc reinterpretation of existing data or on knowledge obtained through successful predictions and observations; (7) consistency of the opinion with accepted theories, that is, general acceptance in the witness's and related communities of specialization; (8) application and use of the knowledge of the witness by the scientific community outside the context of litigation; and (9) peer review and publication of the knowledge base or, where proprietary information is involved, other internal review.

Though not required under *Daubert* or *Kumho Tire*, these and other authoritative criteria for determining whether a field of knowledge is science may be invoked by counsel as a basis for establishing or challenging an area of specialized knowledge that is claimed to be a science. Such knowledge-based inquiries into the validity and reliability of a particular field of specialization are supported (and arguably required) by *Daubert*'s focus on validity and reliability. Challenges to forensic evidence from fields of knowledge outside the witness's field of knowledge may become the rule, rather than the exception, for cross-examination of witnesses in fields of forensic knowledge that have matured under the protection of general acceptance by only their guilded peers.

Assistance to trier of fact. Whether the forensic evidence will assist the trier of fact to understand the evidence or determine a fact in issue has been defined as whether there is a "fit" between the evidence and the issues in the case. Valid, reliable evidence offered by a qualified witness will be inadmissible if it does not assist the trier of fact. The Supreme Court's example of this element makes the point very well.

"A study of the phases of the moon, for example, may provide valid scientific 'knowledge' about whether a certain night was dark, and if darkness is a fact in issue, the knowledge will assist the trier of fact. However (absent creditable grounds supporting such a link), evidence that the moon was full on a certain night will not assist the trier of fact in determining whether an individual was unusually likely to have behaved irrationally on that night. Rule 702's 'helpfulness' standard requires a valid scientific connection to the pertinent inquiry as a precondition to admissibility."

Qualifications. The threshold inquiry into the witness's qualifications may address any and all formal education the witness has or has not completed. Education in the area of specialized knowledge in which the witness claims to be proficient will be most relevant. Any publications or other educational materials authored or edited by the witness may be reviewed and addressed on cross-examination, particularly if they make statements inconsistent with any statements made by the witness in the present case. Professional or technical training in the area of specialization, the witness's performance in that training, and certifications or other credentials related to the areas of specialization may also be addressed.

The witness's experience in the field of specialized knowledge may also be covered, together with any issues related to the witness's competent performance in the field. Performance evaluations, disciplinary proceedings, and any other evidence relevant to the witness's experience may be scrutinized. Testimony of the witness in other court proceedings may also be addressed if it may be used to undermine the reliability or validity of the witness's opinion in the case at hand.

Testimony. Though each deposition and court appearance is a unique experience, forensic witnesses may reasonably expect that direct and cross examination will cover at least the key areas addressed above. Direct examination will progress in a manner that allows the witness to favorably state the witness's education, training, and experience in the witness's field of specialized knowledge; the facts of the case on which the witness relied in preparing any reports, summaries, or formulating any opinions for the case; any theories, techniques, methods, or procedures applied by the witness in understanding and interpreting the facts; and any conclusions or opinions the witness may have reached as a result of this process. Direct examination is typically a smooth, comfortable exposition of the witness's qualifications, reasoning, and opinion.

Cross-examination by opposing counsel is typically more challenging. The primary goal of cross-examination is to identify any weaknesses that may undermine the relevance, reliability, and/or validity of the witness's evidence. Weaknesses will be sought in the witness's qualifications; knowledge of the scope, limits, validity, and reliability of the witness's area of specialization; application of the area of specialized knowledge to the facts of the case; and/or relevance of the evidence to the issues in the case. The forensic witness must, therefore, be prepared for intense scrutiny in any of these areas.

The intensity of cross-examination across the range of areas that may affect validity and reliability of forensic evidence has occasionally been interpreted by some forensic witnesses as a personal attack. Personalities aside, the forensic witness may be better prepared to face intense cross-examination by recalling that the forensic expert is, by definition, present to serve the court, and counsel is present to serve the court and a client. Counsel has a duty to assure that evidence presented against the client is valid and reliable; the forensic witness may request assistance from the court if cross-examination interferes with the witness's ability to serve the court's fact-finding function.

Court-appointed experts. Trial courts have the authority to appoint independent experts to consult with the court concerning issues presented in a case, but they rarely use this authority. *Daubert*'s instruction that the trial judge serve as "gate-keeper" may increase the frequency of appointment of experts by courts. The American Association for the Advancement of Science has launched a program to assist federal judges in identifying appropriate court-appointed experts in civil cases. *See* FORENSIC CHEMISTRY; FORENSIC MEDICINE. Stephen A. Brunette

Bibliography. B. Black and P. Lee (eds.), *Expert Evidence: A Practitioner's Guide to Law, Science, and the FJC Manual*, 1997; J. Cecil and T. Willging, *Court-Appointed Experts: Defining the Role of Experts Appointed Under Federal Rule of Evidence 706*, 1993; D. Faigman et al., *Modern Scientific Evidence: The Law and Science of Expert Testimony*, 1997; Federal Judicial Center, *Reference Manual on Scientific Evidence*, 1994; P. Giannelli and E. Imwinkelried, *Scientific Evidence*, 2d ed., 1993; A. Moennsens et al., *Scientific Evidence in Civil and Criminal Cases*, 4th ed., 1995.

Forensic medicine

The branch of medicine concerned with the resolution of legal issues by the application of scientific medical knowledge. The issues may be of criminal or civil nature. The specialty of forensic medicine is often referred to as legal medicine in Europe and in Spanish-speaking countries. The development of forensic medicine began in Europe in the early nineteenth century. Although forensic medicine has its basis in the specialty of pathology, physicians in other specialties, such as psychiatry, orthopedics, cardiology, and neurology as well as specialists in other disciplines, may be called on to resolve legal questions posed by judges, attorneys, investigators, and hearing boards.

In death investigations, a number of forensic specialists may work together. A forensic odontologist may be called in to identify the deceased person through dental examination, or the perpetrator of bite marks left on the deceased's body. A forensic anthropologist may be needed to identify skeletal remains; a forensic toxicologist for the identification of poisons or drugs; and a criminalist for investigation at the scene and collection of evidence, or for study of trace evidence such as blood stains, hair, paints, and seminal fluid. In cases of equivocal suicides, forensic psychiatrists and psychologists may be called in. Recently the biomechanical engineer has been added to the roster of forensic specialists, to test and study injury patterns to determine how the injury came about. The forensic medical specialist may express an opinion in writing or may be required to testify in person in the courtroom. As with all legal proceedings, the physician witness must be cognizant of issues such as the degree of proof, the chain of custody of specimens or evidence, competency of the witness, and court procedures. *See* FORENSIC ANTHROPOLOGY; FORENSIC EVIDENCE.

Medicolegal investigative system. In many countries other than Canada and the United States, medical schools have departments of forensic or legal medicine, and medical students take courses in forensic medicine. In these countries, in cases of medicolegal disputes, the professor of legal medicine at the local medical school is asked to examine the patient or conduct the forensic autopsy, and to offer opinion to the courts, attorneys, or investigators. In the United States, in earlier times, the autopsy service was provided by a local physician hired by the coroner's office, which handled the investigation and determined the cause of death. In many of the states and larger metropolitan areas, beginning with the Commonwealth of Massachusetts and New York City, this system evolved into the medical examiner's system. The medical examiner's office has jurisdiction either statewide or countywide. A federal medical examiner's office has been established in Washington, DC. The American Board of Pathology sets the standards, establishes the qualifications, and certifies the competency of the forensic pathologists.

Most board-certified forensic pathologists work as medical examiners for a state or county. Their work involves death investigation, conducting forensic autopsy to determine the cause and manner of death, and often testifying in court. Few maintain full-time private consultation service. A well-staffed medical examiner's office has toxicological and forensic laboratories; a team of investigators; consultants in such forensic specialties as odontology, anthropology, and entomology; and medical consultants in radiology and psychiatry.

Clinical forensic medicine. Forensic medical specialists also deal with medicolegal issues involving living persons. Psychiatrists may assist with the evaluation of insanity pleas by defendants; a cardiologist may deal with a worker's claim for disability due to heart disease; and an orthopedic surgeon may evaluate the degree of a patient's disability for a court or for attorneys prior to settlement. In the United Kingdom and former commonwealth countries, and in the Far East, a system of police surgeons is set up to provide medical advice to the local police. A physician appointed as police surgeon is called on to visit the jail to examine a detainee, or is called to the scene of death. In the United States, clinical evaluation of jail inmates is provided by individual practicing physicians, and the setup varies with the locality.

Authority and practice. In the United States, deaths that are sudden and unexpected or violent are usually investigated by the medical examiner's office, except in areas where the coroner still has jurisdiction. Even in those jurisdictions, most autopsies are conducted by board-certified forensic pathologists. Throughout the world, such deaths are investigated by an official agency. In continental Europe and the Far East, the police conduct the investigation, with an autopsy authorized by a magistrate. In the United Kingdom and former commonwealth countries, the coroner, who is also the magistrate, conducts the inquest. The medical examiner in the United States is a board-certified forensic pathologist, who serves as an appointed rather than an elected official. In all states there are state or countywide medical examiner's offices. Their purpose is to determine by investigation the cause and manner of death. Typical cases include gunshot wounds, stab wounds, blunt-force injuries, traffic accidents, drowning, hanging, asphyxiation, and drug and alcohol abuse. Any death suspected of being an accident, suicide, or homicide is investigated.

The medical examiner's authority has been expanded to cover such incidents as death in nursing homes, sudden and unexpected death during surgery, or death of a prisoner in custody. Furthermore, the medical examiner's office usually takes jurisdiction over cases of undiagnosed epidemic diseases that constitute a public hazard. In some jurisdictions, any proposed cremation of the remains is usually subjected to inquiry by the medical examiner's office prior to the cremation.

The medical examiner's office not only certifies the cause and the manner of death, but also helps prevent accidental, suicidal, and homicidal deaths by maintaining and disseminating biomedical information and by participating in professional and public education. *See* PUBLIC HEALTH.

With the advent of transplants as therapeutic procedures, some medical examiners have become involved in the administration of organ and tissue donation programs and related ethical matters. This necessitated a new legal definition of death. Moreover, ethical problems have arisen concerning the patient's right to refuse certain kinds of care. When a patient is dying or irreversibly comatose, such decision-making is a troubling matter. In some states, a surrogate can be authorized to make decisions about the patient's care (power of attorney for health care). Forensic medicine plays a vital role in the enactment of appropriate legislation to handle these and other ethical matters. *See* DEATH.

Drug and poison detection. Drug and alcohol abuse have been social problems for centuries. Before World War II, alcohol, barbiturates, heroin, aspirin, cyanide, and arsenic were the chemicals most commonly causing death. In the 1960s, stimulants and hallucinogenic drugs arrived. In the 1970s, phencyclidine (PCP) and lysergic acid diethylamide (LSD) caused major problems and deaths, and in the 1980s cocaine became the most widely publicized illegal drug involving fatalities. The medical examiner determines the cause and manner of death due to drugs and alcohol, but also assesses drug and alcohol involvement in deaths due to traffic accidents, homicides, and suicides, and maternal and perinatal deaths.

The rapid improvement in the accuracy and sensitivity of analytical instrumentation, coupled with the availability of sophisticated computers, now makes it possible to detect drugs, poisons, and their metabolites in samples as small as 10^{-12} g. Although the incidence of drug overdose leading to deaths has declined, alcohol, multiple therapeutic drugs, and illegal substances continue to plague society. *See* FORENSIC CHEMISTRY; FORENSIC TOXICOLOGY.

Scanning electron microscope. As part of the medicolegal investigation, the medical examiner performs microscopic examination of tissues. The analysis of trace materials, including gunshot residue, fibers, glass, and serological material, may greatly add to the understanding of the circumstances surrounding death. Microscopic examination, in particular with the scanning electron microscope (**Fig. 1**), which permits viewing of minute objects with magnifications up to 100,000 times, is useful for the analysis and identification of minute fragments of objects left on human bone, cartilage, and soft tissue, such as the crushing marks of an assault tool on the skin and serrated markings on bone in the case of stabbing. Another example is electrocution, where the skin may show the burn of the electric wire and a deposit of copper. Scanning electron microscopy coupled with energy-dispersive x-ray spectroscopy can effectively identify chemicals deposited on the wound and may be used to identify gunshot residues, which are composed principally of lead, antimony, and barium. Atomic absorption spectroscopy and neutron activation analysis are also used for this purpose. *See* ANALYTICAL CHEMISTRY; SCANNING ELECTRON MICROSCOPE.

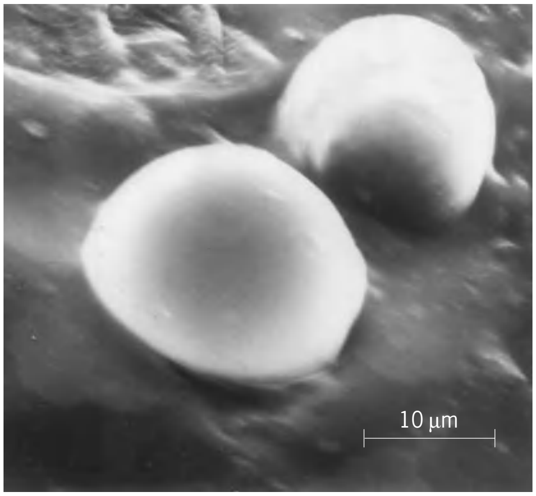


Fig. 1. Scanning electron microscopic image. The two round objects represent molten primer compound emitted from a recently fired cartridge. Energy-dispersive x-ray spectroscopy produced the spectrum characteristic of gunshot residue. (*Office of Chief Medical Examiner-Coroner, County of Los Angeles*)

Identification of bite marks. During the investigation of a death, a forensic pathologist looks for bruises and marking on the victim's skin. The assailant may leave a bite mark on the skin, which can be used to identify the person. Bite marks are second only to fingerprints in their forensic value for identifying murderers, particularly serial killers.

Genetic analysis. Another important development is the forensic application of genetic analysis. Every individual has a unique genetic content determined by deoxyribonucleic acid (DNA) sequences. The DNA present in forensic samples such as hair, bloodstains, and seminal fluid can identify a suspect. One analysis technique is called DNA fingerprinting; it can be used to compare the DNA sequence of a suspect or a defendant with material evidence at the crime scene. The DNA is broken into unique fragments by restriction endonucleases and then separated by electrophoresis. The pattern seen after hybridization with specific probes is called the DNA fingerprint, and like an ordinary fingerprint is

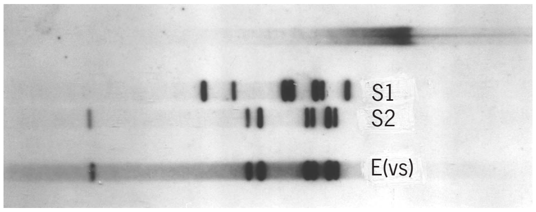


Fig. 2. DNA sequence of two rape-murder suspects, S1 and S2, are compared with the DNA sequence of the seminal fluid found in the vaginal smear from the victim. Suspect sample S2 matches the DNA sequence of the seminal fluid E(vs). A guilty verdict was rendered against suspect S2. (*Cellmark Diagnostics, Division of ICI Americas Inc.*)

characteristic of the individual (**Fig. 2**). DNA analysis is also used in identification of an unknown person or of parts of a human body. Paternity exclusion by ABO blood typing has been replaced by DNA fingerprinting. *See* DEOXYRIBONUCLEIC ACID (DNA).

Thomas T. Noguchi

Bibliography. K. A. Bennett, *A Field Guide for Human Skeletal Identification*, 2d ed., 1993; S. B. Karch, *The Pathology of Drug Abuse*, 3d ed., CRC Press, Boca Raton, 2001; N. Rudin and K. Inman, *An Introduction to Forensic DNA Analysis*, CRC Press, Boca Raton, FL, 2d ed., 2001; W. U. Spitz, *Spitz and Fisher's Medicolegal Investigation of Death: Guidelines for the Application to Crime Investigation*, 3d ed., Charles C. Thomas, Springfield, IL, 1993; C. V. Wetli, R. E. Mittleman, and V. J. Rao, *An Atlas of Forensic Pathology*, ASCP (American Society of Clinical Pathologists) Press, Chicago, 1999.

Forensic microscopy

The application of microscopy techniques for purposes of civil or criminal law. Edmond Locard (1877–1966) stated that every contact leaves a trace. This is known as the Locard exchange principle, the basis for much of forensic microscopy. Since the early 1800s, the microscope has been used to help solve crimes. Today, it remains one of the most used tools in the crime laboratory. Criminal evidence ranging in scale from micrometer-sized particles to hair and paint chips can be found and removed to a laboratory for microscopic analysis, where its history may be deduced to help solve the crime. Microscopy can provide insight into the identity and origin of a material, what has happened to it and when, and the routes it may have taken between a crime victim, suspect, and crime scene (**Fig. 1**). *See* CRIMINALISTICS.

Stereomicroscope. The low-powered stereomicroscope is probably the most used microscope in the crime lab. The magnification typically ranges from 7× to 40×, and the image is upright (not reversed) and three-dimensional. With additional lenses, the magnification can range from below 3× to above 100×. This type of microscope can be used with a variety of lighting techniques such as transmitted, fluorescence, diffuse, coaxial, and oblique illumination. *See* FLUORESCENCE MICROSCOPE; MICROSCOPE; OPTICAL MICROSCOPE.

The stereomicroscope is used in drug analysis to help identify marijuana and to evaluate powders and chunks of material that may need to be separated before analysis. In examining firearms and tool marks, it is used to evaluate striations or imprint marks, look for trace evidence, evaluate bullet holes for muzzle blast, and search for unburned gunpowder particles. In addition, when a bullet passes through an object, some of that object may be transferred to the bullet and detected microscopically. In forensic biology, the stereomicroscope is used to find and measure small bloodstains, to evaluate a bloodstain to determine if it came from the inside or outside of a garment, and to remove stains from hidden places. When examining documents, it can be used to evaluate typefaces, examine strikeovers and ink lines that cross over, and find erasures or alterations. For trace evidence, the stereomicroscope is the tool of choice for finding and characterizing hairs, fibers, paint, glass, soil, and building material. *See* FORENSIC BIOLOGY.

Biological microscope. The biological microscope is used for examining biological fluids, human tissue, hairs, fibers, drug crystals, food, and other stains. A typical biological microscope has 4×, 10×, 20×, and 40× objectives on a revolving nosepiece. It is normally equipped with 10× eyepieces giving a final magnification of from 40× to 400×. Many additional lenses can be added to this style of microscope to increase the range of magnification from less than 20× to greater than 1500×. In addition, these microscopes can be modified to include darkfield, phase contrast, fluorescence, dispersion staining, darkfield epi-illumination, and reflected coaxial illumination. When darkfield epi-illumination or reflected coaxial illumination is used, the biological microscope is transformed into a metallurgical microscope, which

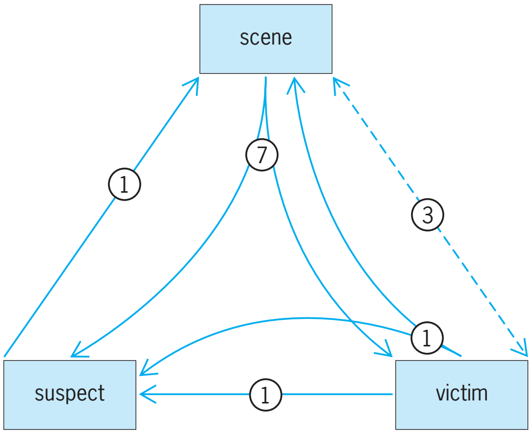


Fig. 1. Hypothetical flow of trace evidence between a crime scene, victim, and suspect. Seven types of trace evidence were transferred from the crime scene to both the suspect and victim. Three types of trace evidence were found at both the crime scene and on the victim. One type of trace evidence was transferred from the victim to the suspect and the crime scene. One type of evidence was transferred from the victim to the suspect. One type of evidence was transferred from the suspect to the crime scene.

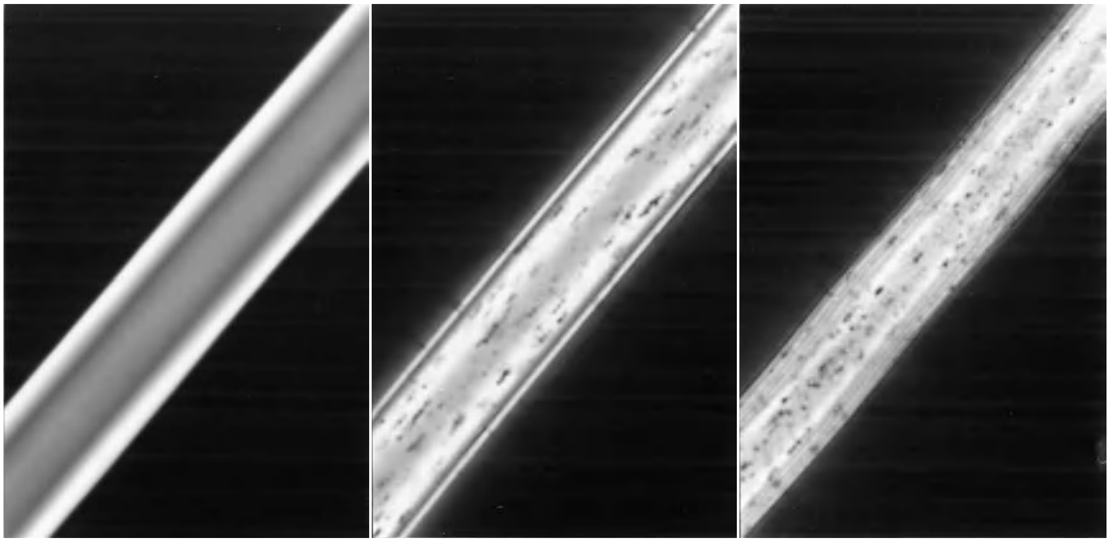


Fig. 2. Crossed polarizers showing interference colors (represented here in black and white) of three colorless synthetic fibers. These interference colors can be used to calculate the birefringence, which is an optical property used to identify these fibers. The birefringence of polypropylene is medium, nylon is high, and polyester is very high.

can be used to examine surface detail and opaque objects. *See* PHASE-CONTRAST MICROSCOPE; REFLECTING MICROSCOPE.

A forensic pathologist typically uses the biological microscope to examine tissue samples from injuries to determine if the injuries occurred before, during, or after death. In the forensic biology section of the laboratory, the microscope is typically used to identify sperm in rape cases. This process involves differential staining and examining the slide at greater than 200×. When sperm are identified from samples collected from a victim, this is proof that sexual activity has taken place. The legal issue of consent is not addressed by the presence of sperm. Other body fluids such as vaginal secretions, saliva, and feces might have been mixed with semen, and microscopy can give vital clues as to the origins of these other body fluids.

Fig. 3. Fifteen anagen hair roots sent for DNA typing.

Some crime labs use crystal tests to identify drugs such as phencyclidine, amphetamines, cocaine, barbiturates, heroin, morphine, and codeine. In a typical crystal test, a small amount of the drug and a liquid reagent are placed on a microscope slide. After a short time, unique crystals for that drug can be recognized at 100× with a biological microscope.

Polarized light microscope. The polarized light microscope is a biological microscope with special modifications, including a polarizer, analyzer, rotating stage, accessory slot, Bertrand lens, and a flip-up condenser. This microscope is typically used to identify synthetic fibers, minerals, glass, starch, and many other types of particles. It has the advantage of measuring optical properties without altering the sample. Six major synthetic fiber types (acrylic, acetate, rayon, olefin, nylon, and polyester) can be quickly differentiated with the polarized light microscope (**Fig. 2**). A significant amount of training and experience, as well as reference standards, is necessary to master the use of this instrument, since most universities do not teach this subject. The McCrone Research Institute in Chicago has trained many forensic scientists in the use of the polarized light microscope. *See* BIREFRINGENCE; POLARIZED LIGHT MICROSCOPE.

Comparison microscope. A comparison microscope consists of two identically equipped compound microscopes joined together with an optical bridge, giving a split-screen image. The comparison microscope is used to compare bullets, cartridge cases, and toolmarks. This microscope normally uses low-powered objectives and reflected, diffused, or oblique illumination. The comparison microscope is also used to compare trace evidence such as hairs and fibers. It uses the same range of magnifications as the biological microscope.

One such case illustrating the comparison of hair, as well as DNA analysis, involved an elderly male victim who was stabbed in a robbery at a hamburger stand. Clumps of long dark hair that appeared to have been pulled out during a struggle were found in several locations at the crime scene. Microscopic examination of that hair showed more than 50 anagen roots. Anagen roots are those still attached to the scalp with a blood vessel and a nerve, indicating that the clumps of hair were forcibly removed during the struggle. The hair from the crime scene had a maximum length of 17 in. The maximum length of the victim's gray hair was less than 1 in. The police collected and the crime lab compared hair from several long-haired suspects with negative results. Finally, a female drug user was identified by a relative as the murderer. A hair sample was collected and submitted to the hair examiner, who testified that the hairs from the crime scene probably came from the suspect. A blood sample from the suspect along with 15 roots from the crime scene hairs were submitted to a DNA laboratory (**Fig. 3**). The DNA matched, and the suspect was convicted. This case became the first California DNA case to successfully get through the Appeals Court.

Scanning electron microscope. The scanning electron microscope with energy dispersive x-ray spectrometer (SEM/EDX) is routinely used to identify the elemental content of very small samples. The most common sample analyzed with the SEM/EDX is gunshot residue. This is done by dabbing the shooter's hands with sticky tape attached to an SEM stub (a substrate for mounting specimens). The stub is then examined for gunshot residue (GSR) particles (usually 1–6 micrometers) that contain lead, barium, and antimony. Many of the SEM/EDX microscopes are connected to computers that have programs that can search automatically for GSR particles. *See* SCANNING ELECTRON MICROSCOPE.

Digital forensic microscopy. Most microscope systems can be modified to capture digital images, which can be sent by e-mail to forensic scientists for examination and comment. An automated system for glass refractive index measurement uses a phase-contrast microscope, hot stage, video system, and computer to compare glass samples. The National Integrated Ballistics Identification Network (NIBIN), sponsored by the Federal Bureau of Investigation (FBI) and Bureau of Alcohol, Tobacco and Firearms (ATF), uses digitally captured images (through the microscope) of cartridge cases or bullets to link shooting incidents with guns taken from suspects in distant jurisdictions.

The infrared spectrophotometer with a microscope attached is a common tool in crime laboratories. It is used to perform spectrophotometric analysis on very small samples such as paints, plastics, fibers, and many other substances. The microspectrophotometer is used to measure the spectrum of the visible colors, ultraviolet radiation, and fluorescence that are present in microscopic objects. An automated version of the microspectrophotometer using a remote control stage and computer can search tape lifts taken from suspects, victims, and crime scenes. First a fiber is chosen and entered into the apparatus. The tape lifts are then searched for that target fiber.

Another digital system is a portable microscope with fiber-optic cables that is used to acquire high-quality microscopic images such as a paint transfer on a bumper or hood of a vehicle without dismantling.

Edwin L. Jones, Jr.

Bibliography. P. De Forest, Foundations of forensic microscopy, in R. Saferstein (ed.), *Forensic Science Handbook*, Prentice Hall, Upper Saddle River, NJ, 2002; B. Fisher, *Techniques of Crime Scene Investigation*, CRC Press, Boca Raton, 1993; J. Houde, *Crime Lab: A Guide for Nonscientists*, Calico Press, Ventura, 1999; S. Palinik, Microscopy, in J. A. Siegal et al. (eds.), *Encyclopedia of Forensic Sciences*, Academic Press, San Diego, 2000; R. Saferstein, *Criminalistics: An Introduction to Forensic Science*, Prentice Hall, Upper Saddle River, NJ, 2001.

Forensic physics

The application of physics for purposes of civil or criminal law. Forensic science has been dominated by the fields of chemistry (mainly analytical chemistry) and biology (serology and DNA profiling). Indirectly, physics has contributed to forensic science via the invention of the microscope, the electron microscope, the mass spectrometer, and optical spectrometers; but directly, its role in forensic science has been minimal. Forensic physics has traditionally involved the measurement of density (soil and glass examination), index of refraction, and birefringence (fiber analysis, glass examination). In the last 25 years, the use of the photoluminescence phenomenon for physical evidence examination has emerged, with latent fingerprint detection the most notable application. In criminalistics, fingerprint detection is important because it provides absolute identity, and does not suffer from the contamination problems to which DNA profiling is prone. *See* CRIMINALISTICS.

Photoluminescence. Light emission by substances can have various causes, such as heat (incandescence), electricity (electroluminescence), or a chemical reaction (chemiluminescence), as in the reaction of luminol (3-aminophthalhydrazide) with blood, a much-used criminalistics procedure. The most important origin of luminescence in forensic science is the prior absorption of light by the substance and subsequent emission of light at longer wavelength, called photoluminescence. The photoluminescence phenomenon was used in various science fields long before its quantum-mechanical underpinnings were understood. Early on, two types of photoluminescence were distinguished on the basis of photoluminescence lifetime, namely fluorescence (of short lifetime) and phosphorescence (of long lifetime). On practical grounds, this distinction suffices. Fluorescence techniques are presently more widely used in criminalistics than phosphorescence techniques, simply because intense fluorescence is generally encountered

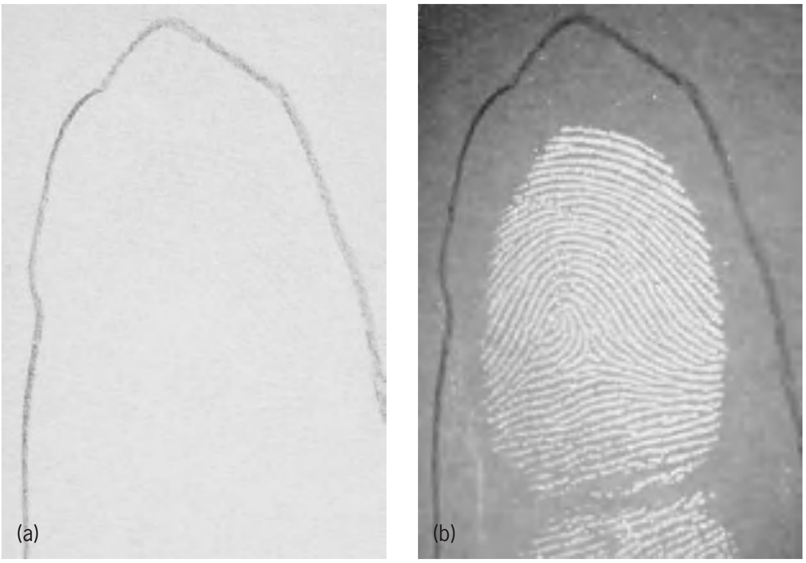


Fig. 1. Fingerprint on paper (*a*) after standard ninhydrin development and (*b*) after subsequent processing by zinc chloride and laser photoluminescence excitation.

in nature more frequently than phosphorescence. Phosphorescence approaches, however, are beginning to show much promise, especially in instances in which fluorescence approaches fail. *See* LUMINESCENCE; PHOTOLUMINESCENCE.

Fluorescence techniques. In the past, a white surface would typically be dusted with a black powder to reveal a latent fingerprint. This mode of visualization involves absorption/reflectance phenomena. In the case of a weak print, to which only few of the absorbing black powder particles adhere, the print would reflect only slightly less than the surroundings. This inherently limits sensitivity because of the small difference between the two relatively large light signals. The photoluminescence approach detects a small signal, rather than a small difference between large signals and is therefore inherently more sensitive. Thus, photoluminescence-based techniques have generally supplanted absorption/reflectance techniques in most scientific fields. The principal application of photoluminescence in forensic science is fingerprint detection. Fingerprint residue left on an article typically has a mass of about a tenth of a milligram, and is composed primarily of water (about 98–99%), which soon evaporates, with the remaining residue made up about equally of inorganic material (mainly salts, largely useless for detection purposes) and a myriad of organic compounds. Thus, only nanograms of organic material are available to be made visible to the naked eye. Not only does the material have to be detected, but also its spatial distribution has to be visualized in fine detail. Thus, lasers, by virtue of their power (at the right wavelength to excite luminescence) were the first choice of illumination for purposes of photoluminescence detection of fingerprints (**Fig. 1**). Filtered lamps have since seen wide use, but sacrifice sensitivity in favor of price. Other photoluminescence applications in criminalistics include fiber analysis (especially in finding otherwise elusive fibers), document examination (detecting altered and erased writing), detection of body fluids (mostly semen and blood), trace evidence detection, authentication of currency and passports, explosives origin specification, and DNA profiling (with the radioactive tags of the past replaced by fluorescent tags). *See* FINGERPRINT; FLUORESCENCE.

Phosphorescence techniques. In fingerprint detection, one often encounters background fluorescence from the substrate on which the fingerprint is located, and this background fluorescence, which typically has a lifetime on the order of a nanosecond, cannot always be eliminated by optical filtering. Detection sensitivity is therefore reduced. This is a general problem with analytical fluorescence methods.

There has emerged in recent years a procedure whereby fingerprints are treated so that they exhibit phosphorescence, and then time-resolved imaging techniques are applied to suppress background fluorescence, to detect otherwise elusive prints. In such techniques, the imaging device is turned on and off in synchronization with the photoluminescence excitation source, which is also periodically turned on and off. Recording by the imaging device is delayed with respect to the excitation source turn-off, such that the background fluorescence has decayed by the time the imaging device turns on. Likewise, the imaging device turns off before the onset of the next illumination pulse (**Fig. 2**). Typically, a charge-coupled device (CCD) camera equipped with a proximity-focused microchannel plate image intensifier is used as the imaging device. The time-resolved detection technique is not yet routinely used in crime laboratories, but operational prototype instruments have been in existence in United States and Canadian university forensic science laboratories since the mid-1990s. *See* CHARGED-COUPLED DEVICES; PHOSPHORESCENCE.

A number of fingerprint treatments that lend themselves to time-resolved imaging, mostly based on lanthanide compounds that display millisecond

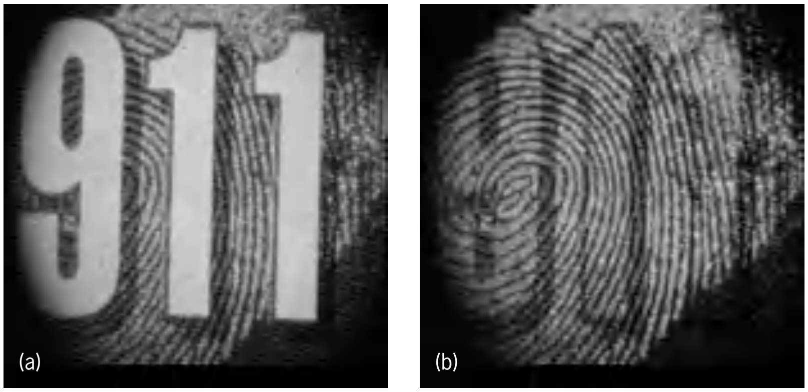


Fig. 2. Fingerprint (*a*) after standard photoluminescence development and (*b*) after time-resolved development. Note the suppression of the background fluorescence from the numbers 911.

luminescence lifetimes, have been developed. Time-resolved imaging techniques typically require phosphorescence lifetimes longer than microseconds. This lifetime requirement places considerable restrictions on the procedures that can be used to treat latent fingerprints.

Treatments that result in fingerprint luminescence lifetimes of 10–1000 nanoseconds can be used if phase-resolved imaging is implemented. Instead of turning the illumination source on and off (by means of a light chopper), the source intensity is modulated sinusoidally. This allows for much higher modulation frequencies than is possible with the abrupt on-off modulation of time-resolved techniques. In response to the sinusoidal luminescence excitation, the luminescence will also be sinusoidally modulated in intensity, but will be phase-shifted by an angle ϕ with respect to the excitation according to $\tan \phi = \omega\tau$, where ω is the modulation angular frequency and τ is the luminescence lifetime. If the gain of the imaging device (CCD camera equipped with a proximity-focused microchannel plate image intensifier) is modulated with the appropriate phase delay versus the modulation of the light source, a suppression of background fluorescence (small ϕ) with respect to the fingerprint luminescence is achieved (optimally ϕ is roughly 45° because phase shifts approaching the maximum of 90° lead to large demodulation of the luminescence intensity). The general scheme here is reminiscent of the lock-in amplifier, long used as a frequency and phase selective device. Phase-resolved imaging has been implemented in cell microscopy, for instance, but instruments have yet to be designed for forensic applications, perhaps because the corresponding fingerprint treatments have yet to reach maturity. There are, however, promising fingerprint treatments in the offing that appear suitable for phase-resolved imaging. The luminescence excitation sources needed for time- and phase-resolved imaging will have to be lasers.

Nanoparticles. Semiconductor material, when fabricated in size on the order of nanometers, may exhibit luminescence intermediate between what would be obtained from the bulk material and would be obtained from individual ions. The photoluminescence lifetimes of such nanoparticles tend to be 10–1000 ns. The color of the luminescence and the luminescence lifetime can be tailored by particle size. The broad absorption spectrum of the particle is typical of semiconductors, and thus amenable to optical luminescence excitation spanning a wide range of wavelengths; whereas the particle's luminescence is sharply defined, akin to molecular fluorescences, and occurs at the (long-wavelength) absorption band edge. Nanoparticles are ideally suited for fingerprint work in terms of their spectroscopic properties, especially in their ability to suppress background fluorescence.

Presently, nanoparticles are used as selective taggants mostly in the biosciences, but the tagging chemistries should apply to fingerprint detection, making this an active area of current forensic research. Forensic application at present focuses on cadmium sulfide nanoparticles incorporated into tree-shaped polymers called dendrimers. The dendrimers have terminal functional chemical groups that selectively tag fingerprints. Cadmium selenide nanocrystals are also used. These crystals are typically encapsulated in zinc sulfide (or silica). The encapsulant protects the nanocrystal, such that its optical properties are not compromised, and functions as the attachment site for molecules (conjugating ligands), which chemically react with the material to be detected (**Fig. 3**). *See* NANOPARTICLES.

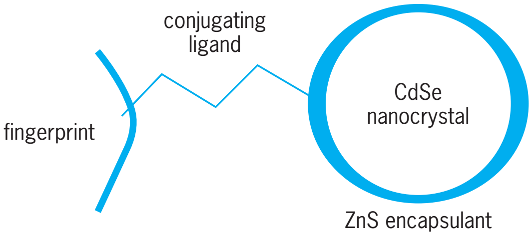


Fig. 3. General scheme for fingerprint detection with nanocrystals.

Given that nanotechnology is a very active field today, with much bioscience application, it would seem to be only a matter of time before it finds its way to routine forensic science work. To date, financial considerations and the conservative nature of the forensic science community itself have been obstacles to its implementation. *See* NANOTECHNOLOGY.

E. Roland Menzel

Bibliography. H. C. Lee and R. E. Gaensslen, *Advances in Fingerprint Technology*, Elsevier, 1991; E. R. Menzel, *Fingerprint Detection with Lasers*, 2d ed., Marcel Dekker, 1999; E. R. Menzel, *Laser Spectroscopy: Techniques and Applications*, Marcel Dekker, 1995; E. R. Menzel, Photoluminescence detection of latent fingerprints with quantum dots for time-resolved imaging, *Fingerprint Whorld*, 26(101):119–123, 2000; E. R. Menzel et al., Photoluminescent CdS/dendrimer nanocomposites for fingerprint detection, *J. Forensic Sci.*, 45(4):770–773, 2000; E. R. Menzel et al., Photoluminescent semiconductor nanocrystals for fingerprint detection, *J. Forensic Sci.*, 45(3):545–551, 2000; R. Saferstein (ed.), *Forensic Science Handbook*, Prentice Hall, 1982; R. Saferstein, *Criminalistics*, 6th ed., Prentice Hall, 1998; D. A. Tomalia et al., A new class of polymers: Starburst-dendritic macromolecules, *Polymer J.*, 17(1):117–132, 1985.

Forensic toxicology

An interdisciplinary science dealing with the adverse effects of drugs and chemicals on various biological systems in a medical-legal context. The forensic toxicologist may work with a medical examiner or coroner in order to determine the role that a particular chemical compound may have played in a death. The forensic toxicologist's activities may also involve assessing emergency room patients, helping to determine suitability of an individual for employment or promotion, screening for performance-altering drugs in athletes, and working with law enforcement agencies, for example, performing tests to determine if a driver operated a motor vehicle under the influence of drugs. The forensic toxicologist is involved not only in the analysis of body fluids and tissue for drugs and poisons but also in the interpretation of the resulting information in a judicial context.

Specimens and samples. Because of the medical-legal context of forensic toxicology, specimens must be maintained under a "chain of custody." Each sample must be accompanied by a paper trail allowing for the tracing of the specimen to its origin. The external portion of the chain of custody should contain information such as the collector, date, time, storage (place and conditions), courier, and time and date of delivery to the laboratory. Once the specimen has been accepted by the toxicology laboratory, an internal chain of custody tracks the persons having access to the samples and the procedures applied to those tissues. Each analysis must then be validated. Validation is an exhaustive accumulation of evidence that a given procedure will produce accurate results.

Sources of samples. Death unattended by a physician or occurring under violent, unusual, or sudden unexplained conditions happens in approximately 20% of the population and requires a thorough medical-legal investigation. The forensic toxicologist provides for the isolation of chemicals from biological samples and the subsequent analysis to determine whether a chemical agent played a role in the cause of death. At autopsy, the forensic pathologist collects postmortem specimens (such as blood, brain, liver, kidney, stomach contents, and bile, and on occasion, lung, bone, hair, fingernails, and fat). Following the extraction of chemicals from tissues, these substances are identified and quantified. The distribution of chemicals in the body provides information about the mode of exposure (for example, ingestion, injection, or inhalation) and the time of exposure relative to death. Of growing importance to the field of forensic toxicology is the analysis of metabolites related to the parent drugs. The forensic toxicologist working in a drug abuse program, emergency room, clinical laboratory, or workplace program is limited to specimens such as blood, urine, and vomitus, which are more easily collected from living patients. Recent publications have explored the use of hair, sweat, and oral fluids as potential forensic specimens.

Analytical techniques. Immunoassays for various drugs and classes of drugs have become commercially available, and equipment has been perfected for automating these procedures. Therefore, immunoassay tests may be used by the forensic toxicologist as the first line of testing. Traditionally, the toxicologist has used a number of isolation techniques such as steam distillation, selective solvent extraction, and microdiffusion, depending on the tissue and the analyte. The use of solid-phase extraction has become commonplace and is applicable to semiautomated procedures. Once the substance has been isolated from the biological specimen, techniques such as chromatography (thin-layer, paper, gas, or liquid), spectrophotometry (colorimetric, ultraviolet, visible, and infrared), and immunoassays can be applied to qualitatively and quantitatively determine the drug. The combination of gas chromatography (and more recently, high-performance liquid chromatography) with mass spectrometry is presently accepted as the "gold standard" of analysis.

Drugs, driving, and traffic safety. One application of forensic toxicology that emerged during the 1990s, concomitant with increased emphasis on traffic safety, is the determination of the role of drugs and alcohol on the ability to operate a motor vehicle. Working with law enforcement agencies, the forensic toxicologist has assumed the vital role of providing scientific confirmation of the conclusions reached by the trained drug recognition expert. Analytical methods have been developed and applied to provide unquestioned identification of drugs and their metabolites as well as dependable quantification. The law enforcement officer may be able to recognize the signs and symptoms of a particular drug; however, the forensic toxicologist can best relate the drug at a given concentration to human psychomotor performance.

Interpretation. Once the analytical information has been collected, the forensic toxicologist must interpret the finding with reference to a cause of death or another set of medical-legal circumstances, such as workplace testing or accusation of driving under the influence of alcohol or drugs. Toxicological data can be obtained from pharmaceutical manufacturers, published literature, or various registries. The greatest challenge facing the forensic toxicologist is the interpretation of cases involving combinations of drugs and chemicals and their complex interactions. *See* FORENSIC CHEMISTRY; FORENSIC MEDICINE; TOXICOLOGY.

Jesse H. Bidanset

Bibliography. R. C. Baselt, *Disposition of Toxic Drugs and Chemicals in Man*, 7th ed., Biochemical Publications, 2004; H. Brandenberger and R. A. A. Maes (eds.), *Analytical Toxicology for Clinical, Forensic and Pharmaceutical Chemists*, W. de Gruyter, Berlin, 1997; J. Chamberlain, *Analysis of Drugs in Biological Fluids*, 2d ed., CRC Press, 1995; A. S. Curry, *Poison Detection in Human Organs*, 4th ed., Charles C. Thomas, 1988; M. J. Ellenhorn et al. (eds.), *Ellenhorn's Medical Toxicology: Diagnosis and Treatment of Human Poisoning*, 2d ed., Williams & Wilkins, 1997; L. R. Goldfrank et al., *Goldfrank's Toxicological Emergencies*, 7th ed., McGraw-Hill, 2002; C. D. Klaassen, *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 6th ed., McGraw-Hill, 2001; B. Levine (ed.), *Principles of Forensic Toxicology*, 2d ed., American Association for Clinical Chemistry, 2003; A. C. Moffat (ed.), *Clarke's Isolation and Identification of Drugs in Pharmaceuticals, Body Fluids and Post-Mortem Material*, 2d ed., Pharmaceutical Press, London, 1986.

Forest and forestry

A plant community consisting predominantly of trees and other woody vegetation, growing closely together, is a forest. Forests cover about one-fourth of the land area on Earth. The trees can be large and densely packed, as they are in the coastal forests of the Pacific Northwest, or they can be relatively small and sparsely scattered, as they are in the dry tropical forests of sub-Saharan Africa.

Forests are complex ecosystems that also include soils and decaying organic matter, fungi and bacteria, herbs and shrubs, vines and lichens, ferns and mosses, insects and spiders, reptiles and amphibians, birds and mammals, and many other organisms. All of these components constitute an intricate web with many interconnections. Thus, a bird may be dependent upon the upper branches of a tree for nesting, while the tree may be dependent upon the fungi surrounding its roots to obtain water and nutrients. *See* FOREST ECOSYSTEM.

Forests have important functions, such as cleansing the air, moderating the climate, filtering water, cycling nutrients, providing habitat, and performing a number of other vital environmental services. They also supply a variety of valuable products ranging from pharmaceuticals and greenery to lumber and paper products. *See* FOREST TIMBER RESOURCES; LUMBER.

Forest types. There are many ways to classify forests, as by (1) location (for example, temperate zone forests, tropical zone forests); (2) ownership (for example, public forests, private forests); (3) age or origin (for example, old-growth forests, second-growth forests, plantation forests); (4) important species (such as Douglas-fir forests, redwood forests); (5) economic and social importance (for example, commercial forests, noncommercial forests, urban forests, wilderness); (6) wood properties (for example, hardwood forests, softwood forests); (7) botanical makeup (for example, broadleaf forests, evergreen forests); or (8) a combination of features (such as moist temperate coniferous forests, dry tropical deciduous forests). The last approach tends to be the most descriptive because it often integrates several dominant characteristics related to climate, geography, and botanical features. Some examples of the major forest types are:

1. Northern coniferous forests which span the cold, northern latitudes of Canada and Europe. They are dominated by vast expanses of spruce, fir, and larch, although species of birch, aspen, and pine often are present following fire and other disturbances.

2. Temperate mixed forests which occupy the eastern United States, southeastern Canada, central Europe, Japan, and East Asia, and parts of the Southern Hemisphere in Chile, Argentina, Australia, and New Zealand. They comprise a variety of broadleaf species (such as oak, hickory, beech, and maple in the Northern Hemisphere, and eucalyptus and southern beech in the Southern Hemisphere); subregions such as the southeastern United States also can have large components of pine or mixed pine-broadleaf forests.

3. Temperate rainforests which are situated along moist, coastal regions of the Pacific Northwest, southern Chile, southeastern Australia, and Tasmania. They are frequently dominated by large conifers (such as redwood, Douglas-fir, spruce, cedar, and hemlock) in the Northern Hemisphere and broadleaf species (such as eucalyptus, acacia, casuarina, and araucaria) in the Southern Hemisphere.

4. Tropical rainforests which are found in the equatorial regions of Central and South America (for example, Costa Rica, Brazil, and Ecuador); on the west coast of Africa (for example, Congo, Ivory Coast, and Nigeria); and Southeast Asia (for example, Thailand, Malaysia, and Indonesia). They are typified by a rich diversity of broadleaf trees (such as teak, mahogany, lauan, and meranti) and attendant vegetation (shrubs, vines, and other flora) arrayed in a multilayered canopy.

5. Dry forests which occur in the southwestern United States, the Mediterranean region, sub-Saharan Africa, and semiarid regions of Mexico, India, and Central and South America. They are characterized by relatively sparse distributions of pine, juniper, oak, olive, acacia, mesquite, and other drought-resistant species growing in scrub woodland, savanna, or chaparral settings.

6. Mountain forests which are characteristic of mountainous regions throughout the world. The vegetation varies greatly depending on location, but tends to follow an elevational gradient from open woodland at lower elevations, to mixed conifer-broadleaf forest at middle elevations, to relatively pure conifer forest at subalpine elevations; there may be a wet side and a dry side to the mountain chain due to orographic weather patterns, with characteristic flora associated with each.

Characteristics. Although forests take a variety of forms, they have several features in common that allow them to develop in their respective environments.

Resilience. Forests generally contain a broad array of species, each of which is well adapted to the environmental conditions of the region. This biodiversity and adaptability help the forests cope with natural (and in some cases human-caused) forces of destruction, including wildfire, windstorms, floods, and pests. This built-in resiliency also allows the periodic extraction of wood and other products without jeopardizing the long-term health and productivity of the ecosystems—provided such harvesting operations are performed with care.

Succession. Forests are dynamic—they are constantly changing at both landscape and smaller scales of resolution. This natural propensity to change and develop over time is called forest succession. For example, a forest recovering from a wildfire may be dominated by colonizing species such as grasses, forbs, sprouting shrubs, and fast-growing pines during the first few years or decades. As the forest ages, more shade-tolerant, longer-lived species such as oak, maple, or hemlock may gradually encroach and assume dominance. During the latter stages of forest succession, small patchlike openings may occur in the canopy that allow yet a third wave of plants to become established. In a natural forest, this sequence of events may be repeated in numerous configurations, resulting in a diverse mosaic of vegetation until the next major disturbance occurs. The interval between major disturbances may be as short as a few decades to as long as several centuries. *See* ECOLOGICAL SUCCESSION.

Under managed forest conditions, the process of succession can be modified to favor one stage of forest development over another. In the southeastern United States, it is possible to accelerate the establishment and growth of commercially important pine species and to retard the encroachment of late-successional broadleaf species through a variety of silvicultural practices.

Environmental influences. Forests have a mitigating influence on the environment. This characteristic not only facilitates their own survival and development but also moderates the surrounding climate. Thus, the air temperature in and near forests is lower in the summer and higher in the winter than over open land. Effects of wind and rainfall are ameliorated by the physical structure of forests, thereby reducing wind speed, wind chill, evaporation, soil erosion, soil compaction, and other parameters. Many animal species and humans seek the protection provided by forests in one capacity or another. There also is evidence that removing forests over large areas can affect regional weather conditions.

Ecological processes and hydrologic cycle. Forests play a vital role in ecological processes. From a global perspective, they help convert carbon dioxide in the atmosphere to oxygen, thereby facilitating life for aerobic organisms.

Forests can also capture, store, convert, and recycle a variety of nutrients such as nitrogen, phosphorus, and sulfur. These abilities are based on complex symbiotic relationships among forest plants, microorganisms, soil particles, and other components of the ecosystem. They allow forests to essentially fertilize themselves, thereby maintaining their inherent productivity. This propensity to handle a variety of chemical substances helps forests deal with several industrial pollutants that might otherwise accumulate in the environment to toxic levels. Thus, emissions of sulfur and nitrous oxides can be effectively scrubbed from the atmosphere, provided concentrations are below thresholds toxic to the forest ecosystems themselves. *See* AIR POLLUTION; FORESTRY, URBAN.

Forests also play a critical role in the hydrologic cycle. The canopy of a forest shields the forest floor and soil from the erosive and compacting effects of falling raindrops. In some cases the canopy acts as a large condenser, trapping moisture from fog and low clouds that would otherwise be lost from the area. The spongelike nature of the forest floor acts as a natural filter, cleansing the water and reducing the likelihood of flash floods and erosion. The relatively large storage capacity of most forest soils allows much of the water to be returned to the atmosphere via transpiration from the foliage. *See* FOREST SOIL.

Finally, forests play a crucial ecological role in the habitat that they provide for countless organisms. The need to maintain the health, productivity, and integrity of forest ecosystems is vital to the welfare of humans, now and in future.

Forestry and forest management. The Society of American Foresters defines forestry as the science, the art, and the practice of managing and using for human benefit the natural resources that occur on and in association with forest lands. Natural resources have traditionally entailed major commodities such as wood, forage, water, wildlife, and recreation. However, the concept of forestry has expanded to encompass consideration of the entire forest ecosystem, ranging from mushrooms to landscapes. The practice of forestry requires in-depth knowledge of the complex biological nature of the forest. It also requires an understanding of geology and soils, climate and weather, fish and wildlife, forest growth and development, and social and economic factors.

Foresters, wildlife managers, park rangers, and other natural resource specialists are trained in biology, physics, chemistry, mathematics, statistics, computer science, communications, economics, and sociology. Specialized areas may involve subjects such as plant identification, wildlife biology, forest ecology, forest physiology and genetics, soil science, forest engineering and surveying, hydrology, fire protection, meteorology, biometrics and modeling, wood technology and utilization, and silviculture. Training in forest pest management, wildlife conservation, watershed management, recreation resource management, forest policy and decision making, and business management is crucial.

By combining formal education with practical training and experience, foresters and other natural resource specialists are able to manage forest resources for a variety of purposes. The forest environment can be protected, rehabilitated, or enhanced depending on the circumstances. Effectively manipulating forest stands and landscapes is a prerequisite to sustaining the flow of products, services, and amenities derived from forest ecosystems. *See* FOREST MANAGEMENT.

Silviculture. Silviculture is the art, science, and practice of controlling the establishment, composition, and growth of a forest. It entails the use of both natural and induced processes to foster forest development. For example, reforestation of a harvested or burned-over area can be accomplished by natural seeding from nearby trees or by planting seedlings. *See* REFORESTATION.

Where planting is done, care must be taken to ensure that the trees are adapted to the site. Seed may be collected from mature trees in the area ahead of time. Alternatively, the seed may come from a seed orchard that contains genetically improved trees derived from the general area to be reforested. In either case, the seed are sown in a nursery or greenhouse and carefully cultivated to produce vigorous planting stock. Survival and early growth of the outplanted seedlings can be enhanced by reducing weed competition and by protecting them from animal pests.

Where natural fires have been excluded from forest ecosystems for long periods of time, it is sometimes necessary to reintroduce fire in order to restore forest health and diversity. Foresters can use prescribed burning as well as carefully contained natural fires to accomplish such objectives. In other instances, low-intensity understory burns can be conducted to stimulate forage production for wild-life and to reduce the risk of catastrophic crown fires. *See* FOREST FIRE.

Several other silvicultural techniques are used to foster stand development and maintain forest health. Periodic thinning of the trees helps maintain stand vigor and promotes rapid growth of the forest. Damage by forest insect pests and diseases often can be prevented or reduced by using direct and indirect control measures. Soil fertility can be enriched by favoring nitrogen-fixing plants that occur naturally in the ecosystem or by applying fertilizers. Trees designated for harvest can be pruned to reduce knots, thereby improving wood quality. Finally, a variety of harvesting practices ranging from clearcutting to selective tree removal can be used to provide wood products, wildlife habitats, stream improvements, recreation areas, and combinations thereof. *See* SILVICULTURE.

Forest ownership and administration. Both public and private entities are involved in forest ownership. Public ownership ranges from small municipal forests to large holdings managed by the state and federal governments. Private ownerships range from individually owned woodlots to large corporate holdings. In the United States, for example, the categories of ownership are nonindustrial private (55%), forest industry (15%), federal government (20%), and state and local government (10%). Ownership can vary dramatically by region. For example, two-thirds of the forest land in the western United States is controlled by the federal government, whereas most of the forest land in the eastern United States is owned by individuals.

The goals of forest owners vary greatly, too. For example, the owner of a small woodlot in New England may manage the land for personal enjoyment and occasional revenue. An agroforester in Africa may be more interested in providing fuelwood, food, and watershed protection for families and livestock. An industrial landowner in Brazil may need to grow short-rotation crops of eucalyptus to supply a local pulpmill. Or, a tree farmer in the Pacific Northwest may produce Christmas trees and sawlogs. National Forests in the Rocky Mountains may serve multiple uses, ranging from outdoor recreation to livestock grazing and timber production. National Parks such as Yellowstone and natural wilderness areas are designed to preserve outstanding geological, biological, and cultural features of the landscape. *See* LANDSCAPE ECOLOGY.

Laws and policies. Regardless of ownership types and goals, the management of forest land in the United States is regulated by numerous laws and policies. Federal agencies must comply with laws such as the National Forest Management Act (1976), the Forest and Rangeland Renewable Resources Planning Act (1974), and the National Environmental Policy Act (1969). Other public and private forest landowners generally must comply with state regulations or guidelines designed to promote sound forest stewardship in their respective regions. In some instances, federal laws may apply to both public and private lands; examples are the Endangered Species Act (1973), Clean Air Act (1968), and Water Pollution Control Act (1972). In other instances, public agencies and private owners may adopt self-imposed standards even more stringent than those required by federal and state law in order to meet special needs. Many other countries have developed or are developing analogous laws, regulations, and guidelines governing their forests.

Sound forest policy is a precursor to the regulatory framework protecting the integrity of forest lands and human communities that depend on them. Policy makers in government, industry, environmental organizations, and the private sector strive to balance the multitude of interests surrounding forest resources. Input from the public as well as resource managers and specialists is a crucial ingredient in the process.

Utilization of forest resources. Forests are often focused on particular uses. For example, plantation forests are generally designed to produce wood and fiber products. Conversely, public forests are increasingly devoted to nonconsumptive purposes such as the preservation of biodiversity, natural conditions, and scenic vistas. However, all forests can provide multiple benefits, including harvestable products, watershed protection, recreation opportunities, wildlife habitat, and ecological services. *See* FOREST GENETICS; FOREST HARVEST AND ENGINEERING.

Sustainable forest management. The 1992 United Nations Conference on Environment and Development—the Earth Summit—brought worldwide attention to the issues of forest management, practices, and government policy. The non-legally-binding “Forest Principles” endorsed by 144 countries at the Earth Summit promote the sustainable management of forestlands, which includes conservation of biodiversity, wildlife, and other important environmental values, and the use of forest products for economic and social development. Following the Earth Summit, both governmental and nongovernmental organizations increasingly focused on forestland uses and forest management practices, recognizing that forestlands must be conserved as forests and that forest management practices must be improved to ensure ecological benefits and sustainability over time.

While sustainable forest management was recognized as the goal, there was no internationally accepted formal or informal definition of it. Groups of countries came together to work on these definitions. As a result, the components or elements of sustainable forest management have been identified at the international level for the European countries (Pan-European Process); the non-European temperate and boreal forest countries (Montreal Process), including the United States; the tropical countries (the International Tropical Timber Organization); the African countries (African Timber Organization); and others. Currently there is no legally binding international convention or treaty governing forests or forest management policy.

Forest management certification. Programs to certify reforestation success have been in practice for a number of years. However, consumers in a number of European countries raised concerns about the source of wood products, and questioned whether the wood was coming from forests managed on a sustainable basis, especially wood from tropical rainforests. This phenomenon, along with growing public concern in other developed countries about the environmental impacts of forest management practices, led to development by nongovernmental environmental organizations and the timber industry of new programs to promote, evaluate, verify, or certify the management of forestlands on a voluntary basis by landowners. While the early focus of certification efforts was on tropical forests in developing countries, increasingly the focus has turned to programs in temperate developed countries, including the United States, European countries, and Canada.

Forest management standard-setting and certification programs have been developed by nongovernmental organizations both internationally and within the United States. The forest products industry in the United States has implemented the Sustainable Forestry Initiative^{sm} (SFI^{sm}) program, which sets out principles and standards for forest management by member companies and licensees of the American Forest & Paper Association. The Forest Stewardship Council (FSC) is an international organization that has established broad principles and guidelines for forest management, to be supplemented by regionally developed standards. The Canadian Standards Association has developed a program for use in Canada, and the United Kingdom has developed the UK Forestry Standard with the intent to be compatible with a range of other certification programs. Not all of these programs mandate certification of forest management practices. Some allow for verification by first, second, and third parties as desired by the landowner.

Several programs include the option of labeling wood products as certified under a specific program. Labeling of wood products is a complicated process and, in many countries, is regulated by the government. Claims made by the producers of wood or the manufacturers of wood products are varied from the Forest Stewardship Council claim of wood from "well-managed forests" to Indonesian claims of wood from "sustainably managed" forests. *See* FOREST MANAGEMENT; WOOD PRODUCTS.

Mary J. Coulombe

Bibliography. J. G. Laarman and R. A. Sedjo, *Global Forests: Issues for Six Billion People*, 1992; D. W. MacCleery, *American Forests: A History of Resiliency and Recovery*, 1993; K. Ramakrishna and G. M. Woodwell (eds.), *World Forests for the Future*, 1993; N. P. Sharma (ed.), *Managing the World's Forests: Looking for Balance Between Conservation and Development*, 1992; G. W. Sharpe et al., *Introduction to Forest and Renewable Resources*, 1995; L. C. Walker, *Forests: A Naturalist's Guide to Trees and Forest Ecology*, 1990; J. Westoby, *Introduction to World Forestry*, 1989; R. A. Young and R. L. Giese (eds.), *Introduction to Forest Science*, 2d ed., 1990.

Forest ecosystem

The entire assemblage of organisms (trees, shrubs, herbs, bacteria, fungi, and animals, including people) together with their environmental substrate (the surrounding air, soil, water, organic debris, and rocks), interacting inside a defined boundary. Forests and woodlands occupy about 38% of the Earth's surface, and they are more productive and have greater biodiversity than other types of terrestrial vegetation. Forests grow in a wide variety of climates, from steamy tropical rainforests to frigid arctic mountain slopes, and from arid interior mountains to windy rain-drenched coastlines. Forests with deciduous and evergreen trees are found in the tropics, in the temperate zone, and in the taiga of high latitudes. Forests are found growing on top of moving glaciers; in areas with little rainfall that survive on impacted fog droplets; and in areas with standing fresh and even salt water. The type of forest in a given place results from a complex of factors, including frequency and type of disturbances, seed sources, soils, slope and aspect, climate, seasonal patterns of rainfall, insects and pathogens, and history of human influence.

Ecosystem concept. Because ecosystem boundaries are arbitrarily set, primarily as a research tool, they can be defined at many scales, from a cup of soil to the entire planet. Often forest ecosystems are studied in watersheds draining to a monitored stream. The structure of forested ecosystems in watersheds is then defined in vertical and horizontal dimensions. Usually the canopy of the tallest trees forms the upper ecosystem boundary, and plants with the deepest roots form the lower boundary. The horizontal structure is usually described by how individual trees, shrubs, herbs, and openings or gaps are distributed. Wildlife ecologists study the relation of stand and landscape patterns to habitat conditions for animals.

Photosynthesis. Photosynthesis is the most important metabolic process on Earth, converting solar energy into the food that fuels the food chain, producing gaseous oxygen (O_2) that animals breathe, and controlling how carbon (C), water, and nutrients are cycled, and even how fast rocks are converted to soil. Photosynthesis is, in effect, the basis of life on Earth. Only autotrophs (green plants and some microbes) can produce chlorophylls, pigments that capture sunlight and convert it into chemical energy. The energy of the Sun's photons excites these pigments in a way that leads to the conversion of carbon dioxide gas (CO_2) from the air and water from the soil to simple sugars (photosynthates). These sugars sustain plant growth, creating more leaves, branches, wood, roots, and seeds. Heterotrophs are organisms, including herbivores, saprophytes, and carnivores, that rely entirely on energy captured by autotrophs. *See* PHOTOSYNTHESIS.

Structure. Woody trees and shrubs are unique in their ability to extend their branches and foliage skyward—more than 100 m (330 ft) in some conifers—and, as a stand of individuals, to capture carbon dioxide and most of the incoming photosynthetically active solar radiation. Only about half of the Sun's light-energy spectrum falls in the photosynthetically active radiation spectrum (400–700 nanometers); some light is reflected back to the atmosphere and some passes through leaves to the ground (infrared light). High rates of photosynthesis require lots of water, and many woody plants have deep and extensive root systems that tap stored ground water between rain storms. Root systems of most plants are greatly extended through a relation between plants and fungi, called mycorrhizal symbiosis. In one type called ectomycorrhizae, the fungus forms a mantle around small rootlets. Fungal mycelia grow into the root's cortical cells and receive plant sugars. These sugars help the fungus to extend mycelia 10–100 cm or more into the surrounding soil. Because the hyphae (branches of the mycelia) can be as small as 5 micrometers, and roots are typically greater than 60 μm, the mycelia extend the effective volume of soil from which plant roots obtain nutrients and water. *See* ROOT (BOTANY).

Nutrients. Photosynthesis and growth require many different mineral elements as constituents of metab-olites, as enzyme cofactors, and for cell walls and membranes. About 95% of plant dry weight is made of carbon, oxygen, and hydrogen. Among the remaining essential elements, phosphorus is used in electron-transfer reactions; nitrogen is part of all amino acids and proteins; sulfur is part of some amino acids; calcium occurs in cell walls and membranes; magnesium is a constituent of chlorophyll; and potassium is necessary for proper functioning of leaf stomates (pores in leaves). Micronutrient elements (iron, copper, zinc, manganese, molybdenum, cobalt, boron, chlorine, and possibly others) are required as well. Most of these elements are taken up by plant roots from the soil. A small portion of what plants require comes from elements dissolved in rain or blown onto plants as dust. Uptake of nitrogen from the atmosphere by nitrogen-fixing bacteria, and the plants with which they form symbiotic relations, are exceptions.

Biomass and production. The biomass of a forest is defined here as the mass of living plants, normally expressed as dry weight per unit area. Confusion can result from variable definitions that may include the mass of other living organisms and the mass of dead organisms (detritus). Living microbes and insects in soil have a significant weight per unit area, but higher animals usually make up far less than 1% of the total mass. Forests usually have a lot of nonliving biomass, such as soil humus and litter and logs on the ground, and most of the heartwood of large trees is nonliving tissue. Thus, biomass needs to be carefully defined.

Biomass production is the rate at which biomass is accrued per unit area over a fixed interval, usually one year. If the forest is used to grow timber crops, production measures focus on the biomass or volume of commercial trees (one common measure of volume is the board foot (12 × 12 × 1 inches, or 144 cubic inches). Likewise, if wildlife populations are the focus of management, managers may choose to measure biomass or numbers of individual animals. Ecologists interested in the general responses of forest ecosystems, however, try to measure net primary production (Npp), usually expressed as gross primary production (Gpp) minus the respiration of autotrophs (Ra), where Ra can be thought of as the cost of doing business:

$$\mathrm{Npp} = \mathrm{Gpp} - \mathrm{Ra}$$

The Gpp is extremely difficult to measure directly, although eddy correlation gas diffusion techniques are being tried in large towers on a few forests. The most common approach is to measure cumulative accrual of organic matter in all plants per unit area per year. This approach is hindered by the difficulty in measuring plant organic matter accruals belowground. Another response commonly of interest is net ecosystem production (NEP) usually expressed as

$$\mathrm{NEP} = \mathrm{Gpp} - (\mathrm{Ra} + \mathrm{Rh}) = \mathrm{Npp} - \mathrm{Rh}$$

where Rh is respiration of heterotrophs. Pragmatic strategies include measuring changes in soil organic matter and soil respiration, which includes Rh and root Ra. *See* BIOMASS.

Forest productivity. Productivity is the change in production over multiple years. Monitoring productivity is especially important in managed forests to ensure that practices are not undermining the capacity of the forest soil to regenerate and grow subsequent timber crops or to provide animal habitat or other desired goods and services from the forest through time. Changes in forest productivity can be detected only over very long periods. In addition to annual fluctuations in climate, most forest stands take several hundred years to develop through successional stages, and meaningful changes in productivity may occur only after several successional or management cycles. Concerns over long-term productivity, therefore, mostly focus on changes in the status of soil nutrients, soil organic matter, structure, and other factors thought to control biomass production. *See* BIOLOGICAL PRODUCTIVITY.

Cycles. Forested ecosystems have great effect on the cycling of carbon, water, and nutrients, and these effects are important in understanding long-term productivity. Cycling of carbon, oxygen, and hydrogen are dominated by photosynthesis, respiration, and decomposition, but they are also affected by other processes. Forests control the hydrologic cycle in important ways. Rain falls on forests and drips through the canopy or along tree stems to the soil. Wind-driven fog-water droplets can also be captured by finely divided leaves and needles.

Photosynthesis requires much more water than is required in its products, called photosynthates. Water is lost back to the atmosphere (transpiration) simply to maintain stomates open and to transport nutrients from the soil through the xylem. Water on leaf and branch surfaces also evaporates under warm and windy conditions. Water not taken up or evaporated flows into the soil and eventually appears in streams, rivers, and oceans where it can be reevaporated and moved back over land, completing the cycle.

Forest plants and animals alter soil characteristics, for example, by adding organic matter, which generally increases the rate at which water infiltrates and is retained. Nutrient elements cycle differently from water and from each other.

Rocks already contain most elements required for plant growth, except nitrogen. Elements such as phosphorus, calcium, and magnesium are released from primary minerals in rocks through chemical weathering. Plants can speed weathering, and as a result release nutrients that support plant growth. Elements incorporated into biomass are returned to the soil with litterfall and root death; these elements become part of soil organic matter and are mineralized by decomposers or become a component of secondary minerals. Some elements are held on exchange sites where they can be taken up by subsequent generations of plants. Some elements, such as potassium, calcium, and magnesium, are readily dissolved in water and can move to streams in soil solutions. All elements can leave ecosystems through erosion of particles and then be transported to the oceans and deposited as sediment. Deeply buried sediments undergo intense pressure and heat that reforms primary minerals. Volcanoes and plate tectonic movements eventually distribute these new minerals back to land.

Nitrogen differs by being rare in rocks but the most common gas (molecular N_2) in the atmosphere (78% by volume). Only certain bacteria can form a special enzyme (nitrogenase) which, with substantial consumption of carbohydrates, breaks apart N_2 and combines with photosynthates to form amino acids and proteins. In nature, free-living N_2-fixing microbes and a few plants that can harbor N_2-fixing bacteria in root nodules (called symbiotic N_2-fixing plants, such as legumes and alders) play important roles controlling the long-term productivity of forests limited by nitrogen supply. Because nitrate (NO_3^-) is an easily soluble anion, it moves freely in soil and can be lost by leaching when plant roots are not active. Bacteria that convert ammonium ion (NH_4^+) to NO_3^- (nitrifiers) and bacteria that convert NO_3^- back to N_2 (denitrifiers) are important in nitrogen cycling as well. *See* FOREST SOIL; HYDROLOGY; NITROGEN CYCLE.

Forest disturbance and succession. Changes in the plant species of a forest over 10 to 100 years or more are referred to as succession. Changes in forest structure are called stand development; changes in composition, structure, and function are called ecosystem development. Simplified models of succession and development have been created and largely abandoned because the inherent complexity of the interacting forces makes model predictions inaccurate.

An example model is that the greatest changes occur immediately after the death of the dominant trees through disturbances such as fire, windthrow, flooding, insects, disease, or harvesting. A typical sequence would be: After disturbance, a reorganization phase begins with pioneer species such as herbs, shrubs, and rapidly growing hardwoods, usually along with some plants or sprouts from the previous stand becoming established. This phase is followed by a stem-exclusion or aggradation phase where canopy dominance shifts to trees that grow up through the pioneers' canopy, initially in partial shade. Especially in conifer forests, as these trees dominate and grow, the pioneer species are overtopped and shaded out. Over time, a transition or understory reinitiation phase develops where individual trees die from windthrow, ground fires, or other causes, creating gaps. When gaps between living trees are big enough to allow enough light through, new trees begin to form a secondary, lower canopy. In the so-called old-growth or steady-state phase, stands tend to have scattered large-diameter, old trees forming an upper canopy and one or more lower canopies of younger trees. Standing snags and rotting logs on the ground are also found in most old-growth stands.

Many problems with this general model have been noted. Most important is that disturbances tend to be variable in their nature, extent, and frequency. For example, individual windthrows, ground and crown fires, and harvesting disturbances create nearly unique conditions and result in different species assemblages and growth trajectories. Further confounding the problem, herbivores and diseases may alter species composition, seeds of important species may or may not be present, and disturbances vary unpredictably through time. The result is a complex, and likely unreproducible, set of conditions at any one time. *See* ECOLOGICAL SUCCESSION; FOREST FIRE.

Environmental issues. Forests provide people with many goods and services. Globally, wood is mostly used for heating and cooking. Wood is the dominant structural material for housing and fiber for paper. Alternatives to wood are gaining wider use, often with much greater environmental consequences. Fossil fuel—used in heating and cooking and in the manufacture of concrete, brick, and steel—can replace wood, but it is not a renewable resource. Mining and manufacturing of alternative materials have greater air and pollution side effects than do the forest operations producing a similar quantity of material. Recycling of paper products has greatly helped to reduce demand for raw materials. Clean water is becoming a dominant forest resource in many places, and the forest's role in cleaning pollutants out of air and water is becoming more widely recognized. As people increasingly populate Earth, forests are also becoming a last refuge to experience nature through hunting and gathering, hiking, fishing, nature appreciation, and spiritual renewal. Many people in cities often support protecting forests for their and their children's future use. Concern over biodiversity currently drives legal strategies to limit extractive uses in the United States and elsewhere.

Ecosystem-based management. Forest managers on United States public lands and around the world have struggled to meet competing and changing societal values and needs. Before 1980, many public forests were systematically converted into managed plantations with the primary goal of producing timber, but not without concerns about other resource values, where possible. As the importance of nontimber values gained public support, intense conflicts emerged. Many public agencies responded by adopting loosely defined concepts of ecosystem management. At its core is the idea that ecosystems are complex assemblages of organisms interacting with their environment and changing in complex ways over time, and that science-based knowledge of how ecosystems work is important to maintaining their biodiversity and long-term productivity. The first step in implementing ecosystem management has often been to reallocate or rezone forests to meet new primary objectives. Concepts of joint production, however, are emerging that attempt to find ways to simultaneously manage for multiple objectives by determining compatible groupings of objectives, where possible.

A complementary concept, disturbance-ecology-based management, is also being tried. This strategy centers on the premise that organisms are more adapted to the historical disturbance patterns than to specific successional states, and that management could more closely emulate natural disturbances and ecosystem responses to disturbance as a way to maintain diversity and long-term productivity, and at the same time permit limited resource extraction. The concept is to manage for a distribution of disturbance types, intensities, and frequencies rather than for static successional stages. To the extent that management-induced disturbances begin to mimic natural disturbances, concerns for maintaining biodiversity (especially of soil species unknown at present) and for maintaining important ecosystem processes can be reduced.

Likewise, globally, loosely defined concepts of sustainable development and ecosystem sustainability have gained acceptance. The principal idea is that although forests should be managed for the welfare of local peoples, management should not preclude similar opportunities for future generations. The interpretation of these general concepts as applied to local situations has frequently proved problematic because people's demands are growing and understanding of ecological limits remains inadequate. *See* GREENHOUSE EFFECT. Bernard T. Bormann

Streams

One of the products of an undisturbed forest is water of high quality flowing in streams, which are a biologically diverse and productive ecosystem. The ecological integrity of the stream is a reflection of the forested watershed that it drains. When the forest is disturbed (for example, by cutting or fire), the stream ecosystem will also be altered.

Streams are dependent upon the surrounding forested watershed in at least six ways: (1) The nature and extent of vegetation on the watershed alters the amount of water flowing in the channel and the timing and intensity of floods. (2) Uptake of nutrients by the forest vegetation and soils alters the chemistry of stream water and influences the concentrations of nutrients (for example, ammonium, nitrates, and phosphates essential for life). (3) The presence of vegetation alters the amount of light reaching the stream, influencing both stream temperature regimes and algal productivity. (4) The amount of sediment entering the channel reflects watershed activities. (5) Dead wood falling in the stream alters channel form and provides essential habitat for stream fauna. (6) Organic matter produced in the forest falls into the stream and provides a food resource for stream fauna.

Detritus in stream food webs. Primary productivity in streams flowing through forests is often limited by light because of the intense shade provided by the forest. Because of limited availability of organic matter produced within the stream by algae, stream food webs are also dependent on detritus (that is, nonliving organic matter) which comes from the surrounding forest. Detritus enters the food web both as particles (for example, falling leaves) and as substances dissolved in water. The two sources of detritus have somewhat different fates in the stream food web.

When a leaf falls into a stream, its water-soluble constituents are initially leached. It is then colonized by fungi and bacteria that break down the structural matrix. After the leaf has been partly decomposed, it is the preferred food for a group of aquatic insect larvae called shredders because they tear apart the leaves in the process of feeding on them. Shredders are taxonomically diverse and include stoneflies, caddisflies, and craneflies as well as other groups. The feces of shredders are fine particles of organic matter that combine with fine particulate organic matter from other sources and are further colonized by bacteria. Shredders are extremely important in accelerating the breakdown of leaves in streams and as a source of fine particulate organic matter. When shredders are experimentally removed from the stream, leaf decay rates are less and the amount of fine particulate organic matter generated decreases dramatically.

Some of the fine particulate organic matter generated by shredders is transported downstream in the water column (called seston). It serves as the food source for another group of aquatic insects called filtering collectors. These include insects such as caddisflies that spin silken nets to capture the drifting organic matter as well as any insects that may have been caught in the current. The fine particulate organic matter that remains on the bottom of the stream provides food for a group of insects called collector-gatherers. This group includes taxa like midges and mayflies.

In addition to providing leaves, the surrounding forest is a source of dissolved organic matter to the streams. Dissolved organic matter includes compounds such as dissolved sugars and amino acids, although it is dominated by larger, less biologically available molecules such as humic substances. Dissolved organic matter is material that is leached out of the soil and carried by water to the channel. In the stream, it is taken up by bacteria or is chemically bound to fine particles. Hence some dissolved organic matter becomes fine particulate organic matter and is consumed by the organisms described above. Many bacteria in streams are attached to hard substrates such as rocks or wood and form an active biofilm that contains both algae and bacteria, and extracellular material that cements the entire community on the rock surface. Very small consumers such as protozoa, rotifers, nematodes, and copepods are also a part of the biofilm community. A group of aquatic insects called scrapers feed on this biofilm. Common scrapers include mayflies, beetles, and snails. *See* HUMUS.

The food web in a forest stream is complex. It is supported by small amounts of algae produced in the stream and larger amounts of dissolved and particulate organic matter produced in the forest. The various groups of aquatic insects described above provide the food for predators, which include other aquatic insects such as stoneflies, true bugs, and helgrammites. Vertebrate predators such as salamanders and fishes are also important predators in stream food webs. *See* FOOD WEB.

Woody debris. The surrounding forest contributes structure to the channel in the form of fallen wood that forms debris dams. These debris dams slow the inevitable downstream movement of organic matter, retaining it in the stream so that the fauna can use it as a food resource. Debris dams also create pools in the channel that provide a critical habitat for stream organisms such as fish. Streams in the western United States show increases in stocks of salmon associated with restoring debris dams to forested streams from which wood had previously been removed. In sand-bottom streams, woody debris offers the only available stable substrate and is heavily colonized by aquatic insects.

Flowing water is the most visible component of a stream. Yet there is flowing water that cannot be seen beneath the surface of the sediments, and there is a community of organisms living deep within the sediments. That environment is called the hyporheic zone of (literally, "beneath the flow"). There is exchange of water, nutrients, and fauna between the hyporheic and surface flowing waters. In addition, much of the water entering the stream channel from the watershed flows through the hyporheic zone and, hence, processes occurring in the hyporheic zone influence what is observed in surface stream water.

Watershed disturbance. There are tight linkages between forests and the streams flowing through them. Thus, streams are susceptible to disturbances occurring in the surrounding forest, particularly if the disturbance is in the part of the watershed immediately adjacent to the stream channel (the riparian zone). Forest streams are altered by any practices (for example, lumbering) or chemical input (such as acid deposition) that alter forest vegetation, by the introduction of exotic species, and by the construction of roads that increase sediment delivery to streams. *See* ACID RAIN; STREAM POLLUTION; STREAM TRANSPORT AND DEPOSITION. Judy L. Meyer

Vertebrates

Forest animals are the consumers in forest ecosystems. They influence the flow of energy and cycling of nutrients through systems, as well as the structure and composition of forests, through their feeding behavior and the disturbances that they create. In turn, their abundance and diversity is influenced by the structure and composition of the forest and the intensity, frequency, size, and pattern of disturbances that occur in forests. Forest vertebrates make up less than 1% of the biomass in most forests, yet they can play important functional roles in forest systems.

Energy flow. In forests, energy used by animals usually is acquired by energy flow through decomposers, because wood that dominates the biomass of forests is largely indigestible by most animals. Fungi and some invertebrates can digest the complex molecules of cellulose and lignins that constitute wood. These organisms provide the basis for food for many forest vertebrates (a detrital-based system). The animals that eat the fungi (fungivores) or invertebrates (insectivores) become food for other predators through a trophic system. Animals use the energy that they gain for survival, movement, reproduction, and in the case of birds and mammals, for maintaining a constant body temperature. As energy is used, it is lost as heat. Consequently, proceeding from one trophic level to the next involves a release of heat and a net loss in the energy available to the next trophic level. Because mass and energy are generally interchangeable, there will be less mass produced per unit of time of high-level carnivores in a system than in the herbivores or lower-level carnivores that they eat.

If a disturbance such as a severe fire occurs in a forest, the early stages of succession may be dominated by foliage rather than wood. At this time, the predominant flow of energy is through the herbivores to the carnivores (a grazing-based system) rather than through the decomposers to the carnivores, unless large amounts of dead wood persist after the fire. As succession proceeds, detrital-based systems again predominate until the next disturbance. Late in ecological succession, tree-fall gaps provide small-scale disturbances within which both detrital and grazing-based systems can operate. However, both detrital- and grazing-based systems operate to some degree in all stages of ecological succession. *See* ECOLOGICAL ENERGETICS; TROPHIC ECOLOGY.

Nutrient cycling. In some situations, animals can greatly influence the cycling of nutrients within systems as well as the transport of nutrients among systems. For instance, animals that burrow in the soil (for example, moles) mix nutrients throughout the soil layers and influence water and air flow through the soil. Birds that feed in one system, such as agricultural lands, and roost communally in forests (for example, blackbirds) can deposit great amounts of nitrogen to the forest floor through their feces. Although these dramatic examples of nutrient movement within and among patches are rare, some level of nutrient movement occurs at all times.

Animal habitat. Each animal species has its own habitat requirements, and its own geographic range which is structured by global climate patterns and soil types, which influence forest vegetation patterns and subsequently food and cover resources for each species. Some species occur over continents (for example, white-tailed deer), but others live only within a few counties (for example, the del Norte salamander). Within a geographic range, each individual within a population occupies a home range area within which it can secure the resources that it needs to survive. Because mass and energy are interchangeable, animals of large mass generally have larger home ranges than smaller animals because large animals require more food for survival. Carnivores have larger home ranges than herbivores because carnivores must have a large enough supply of herbivores to eat if they are to survive, and because there is decreasing energy available from one trophic level to another. Within the home range of each animal, food and cover resources are usually distributed within patches, so that an animal must move among these patches to secure these resources. Patches can occur horizontally throughout a home range as well as vertically within a forest. Forest-dwelling vertebrates have home ranges that usually are volumes rather than areas. Indeed, it is the vertical distribution and use of resources that can influence the diversity of animals that can occur in a forest.

Plant-animal interactions. As animals move among patches, they take resources with them from one patch to another. For instance, some plants produce fruits or flowers that are attractive to animals (for example, cedar waxwings and hummingbirds), while others may produce seeds that adhere to the fur or feathers of an animal. The animals move the seeds or pollen from patch to patch, allowing plants to disperse to new areas. Similarly, some species of voles eat the fruits of mycorrhizal fungi, below the soil surface. The voles disperse the ingested fungal spores by depositing their feces in their underground runways. This allows the fungi to spread to new areas. These species of fungi grow on the roots of some trees and shrubs, thereby enabling the plants to be more efficient in acquiring nutrients from the soil. Not all plants rely on animals for dispersal, however; some seeds disperse from wind or water. *See* POPULATION DISPERSAL.

Although animals may influence forest structure and composition through their movement of nutrients and plant propagules, they also can play a role as a form of forest disturbance. Herbivores that select particular plant species for food can nearly eliminate these plant species from systems that the plant would dominate if the herbivores were not present. In the eastern United States, the grazing habits of white-trailed deer can greatly influence the structure and composition of the deciduous forests. Another animal that alters the structure and function of a forest system is the beaver. Beavers change plant species composition by selectively feeding on certain plant species and by altering water levels of forest streams. In addition, they change flowing streams into ponds which cause sediments to accumulate upstream. Continual sediment accumulation causes the pond to fill, eventually forming meadow. As a result, a different animal community will occupy a beaver-influenced forest stream system. Species such as beaver that constitute a small part of the biomass of the system but greatly influence its structure and function are called keystone species.

Some plant species have defense mechanisms against herbivory by animals, and so can persist in a forest. Spines and thorns are physical deterrents against some large mammals. Other plants produce chemicals in their foliage that either can be distasteful (such as saponins) or can decrease the availability of certain ingested nutrients to the herbivores (such as phenols). However, black-tailed deer have a protein in their saliva that can bind with the phenols, inactivating the chemical and allowing access to the nutrients in the plant. Other animals may not contain this protein. The production of phenols in plant foliage in the Pacific Northwest seems generally to increase with increasing amounts of light available to the plants, and therefore can be influenced by disturbance in the forest. *See* HERBIVORY; PLANT-ANIMAL INTERACTIONS.

Forest disturbance. Physical disturbances to forests, such as fire and wind, can influence the structure and composition of the animal community in the system. An animal community is a collection of populations of species that occur in an area. The size, frequency, severity, and pattern of disturbances in a forest can influence each species in a forest differently. If disturbances are small, are infrequent, cause little change, and are widely dispersed (such as an individual tree fall), the effect of each disturbance on a species with a large home range (for example, deer) is probably minimal. However, that same disturbance may greatly influence resource availability for a species with a small home range, such as a salamander that may use the fallen tree as cover. The accumulation of many tree-fall gaps throughout the home range of a species such as an elk can improve food and cover availability for that species.

Frequent disturbances reduce the opportunity for recovery of the system and hence the predominance of grazing- rather than detrital-based energy flow through the system. Disturbance severity can also influence the availability of dead wood or shrubs persisting after the disturbance, and hence opportunities for food or cover for various species. For example, dead trees (snags) that persist following a fire are used by cavity-nesting birds and mammals which would not inhabit that area otherwise.

Global patterns of environmental stability and disturbance frequency seem to influence the diversity of animals. The most diverse animal communities occur in moist tropical forests where severe disturbances are very infrequent, plant species diversity is very high, and forests are tall and provide a large volume of habitat for different species to occupy. Such environmentally stable systems provide long periods of time (many generations) for plants and animals to coevolve. Moving toward the Poles and from hardwood to conifer forests, environmental stability declines (for example, stronger seasonal fluctuations), disturbance severity, frequency, and size may increase (for example, glaciation or fires), and plant species diversity is reduced. There generally is a decrease in animal diversity from the Equator to the Poles, dependent on forest site productivity and plant diversity.

Forest management. Humans use forests for a variety of reasons and so can alter habitat quality for some animal species, often by altering the characteristics of forest disturbance. Disturbances imposed by humans that are more frequent or severe than those which occurred in the past and to which animals are adapted, can lead to decline and endangerment of species in the area. For example, even-aged forest management that completely removes all standing living and dead trees usually is more severe than a disturbance caused by fire or wind. The living and dead trees that remain following a fire or wind disturbance are used by a wide variety of species that otherwise would be absent. Management that retains some living and dead trees may allow some of these species to persist. Similarly, very frequent cutting of forests may not allow the forest to recover to a point that provides some of the large dead trees and logs and fungi necessary for some species to occur. A less frequent harvest may be preferable for these species. Finally, deforestation for agricultural purposes can cause a much more severe impact on species persistence within a region than forest management where stands are allowed to regrow.

Forests managed for timber production, if consideration is given to severity, frequency, size, and pattern of cutting, may allow each species some habitat in a region at all times and opportunities for movement among patches as some stands are cut and others are regrown. Practices that do not provide sufficient patch sizes (home ranges or more), or that are isolated from one another for long periods, such as conversion of forests to urban or agricultural areas, can result in decline or elimination of species from large areas and replacement of these species with others adapted to the new habitat that is created. All forests, fields, and cities are habitats for some species; however, it is the amount and distribution of each of these that will allow species to persist over time and continue to function as a part of these ecosystems. *See* FOREST MANAGEMENT.

Invertebrates

Invertebrates are major components of forest ecosystems, representing most of the biological diversity and affecting virtually all forest processes and uses. Many species are recognized as important pollinators and seed dispersers that ensure plant reproduction. However, some invertebrates that feed on plants and decaying organic matter have been viewed as detrimental to tree growth or commercial production of forest products. Nevertheless, there are indications that even so-called pests may be instrumental in maintaining ecosystem processes critical to soil fertility, plant productivity, and forest health. Consideration of invertebrate diversity and the variety of roles that invertebrates play in forest ecosystems is essential to balanced assessments of their effects and management of forest resources.

Roles. Invertebrates affect forests primarily through the processes of herbivory and decomposition. They are also involved in the regulation of plant growth, survival, and reproduction; forest diversity; and nutrient cycling. Typically, invertebrate effects on ecosystem structure and function are modest compared to the more conspicuous effects of plants and fungi. However, invertebrates can have effects disproportionate to their numbers or biomass. Litter-feeding invertebrates are instrumental in fragmenting organic matter and infusing it with fungi and microorganisms that hasten its decay. As few as 10,000 southern pine beetles (a tiny phloem-feeding insect) can kill an apparently healthy mature pine tree; some seed predators can destroy virtually the entire reproductive effort of a particular plant species.

Changes in population size also affect the ecological roles of invertebrates. For example, small populations of invertebrates that feed on plants may maintain low rates of foliage turnover and nutrient cycling, with little effect on plant growth or survival, whereas large populations can defoliate entire trees, alter forest structure, and contribute a large amount of plant material and nutrients to the forest floor. Different life stages also may represent different roles. Immature butterflies and moths are defoliators, whereas the adults often are important pollinators.

Populations of a particular invertebrate species fluctuate as environmental conditions vary seasonally or annually. Typically, invertebrate populations are small (especially in diverse forests), but some species can increase rapidly under favorable conditions and affect forest structure and function dramatically.

Climatic conditions. Herbivorous invertebrates are affected by climatic conditions and by changes in plant abundance or nutritional quality. All plants produce various physical and chemical defenses (such as tough leaves and toxic compounds) that deter or reduce feeding by herbivores overall, but volatile compounds often attract adapted herbivore species. Differences in chemical composition distinguish plant species and are responsible for preferences by herbivores adapted to these defenses. Chemical composition can vary in a particularly plant as a result of changes in soil moisture or fertility or as a result of injury or crowding. Stressed plants tend to reduce production of defensive chemicals in order to meet more immediate metabolic needs, thereby becoming more susceptible to herbivores. Densely spaced plants of a particular species also are more vulnerable to increased populations of herbivores because of the ease of moving from one host to the next.

Most soil and litter invertebrates depend on warm, moist conditions for survival and may virtually disappear during dry seasons or years, or as a result of canopy removal and exposure of the forest floor to solar heating. Because dead organic matter often contains plant defensive compounds and usually is nutritionally poor (especially in nitrogen), detritivores often have developed mutualistic associations with microorganisms that enhance the nutritional quality of litter resources.

Plant growth, survival, and reproduction. Elevated populations of plant-eating invertebrates can reduce plant growth, survival, and reproduction dramatically. Reduced growth can result from the conspicuous loss of foliage to defoliating species, including moths or sawflies; from less conspicuous loss of buds, young shoots, or roots to budworms or shoot and root borers; and from the loss of soluble photosynthetic compounds to sap-sucking species, such as aphids, spider mites, and root nematodes. Some plant-eating invertebrates are also vectors for plant pathogens, including viruses and bacteria transmitted by sap-sucking invertebrates, and pathogenic fungi and nematodes transmitted by bark- and wood-boring beetles. Severe plant injury or disease often results in death of the host plant. In sufficient numbers several bark beetle species are capable of killing relatively healthy conifers by introducing pathogenic fungi to exhaust tree defenses.

Plant reproduction is affected by invertebrates involved in pollination, seed dispersal, or seed predation. Insects are primary pollinators and seed dispersers for many trees and understory plants, especially in the tropics, where a greater diversity of plants requires more efficient pollen and seed transport than wind. Many ant species are primarily seed predators whose moist, nutrient-rich storage chambers provide ideal locations for seed germination. Seedling growth on abandoned ant colonies typically is more rapid than on surrounding areas. However, seed predation can often destroy the entire seed crop for a particular plant species.

Although the ability of some plant-eating invertebrates to interfere with production of forest products has been the basis for past suppression efforts in forests, growth loss and death of plants are not inevitable results of feeding by invertebrates. Surviving plants often show enhanced growth. Defoliated trees may grow faster for 20–30 years following defoliation than do nondefoliated trees.

This compensatory growth may reflect improved water or nutrient conditions. For example, improved water balance and increased nutrient availability enhance plant survival during drought. Canopy opening increases precipitation penetration through the canopy and reduces evapotranspiration. In some instances, the amount of water reaching the forest floor during the dry season doubles at 20% defoliation, compared to nondefoliated trees. Foliage loss also reduces water use by drought-stressed trees. Whereas desert plants often shed foliage at the onset of dry conditions, water-stressed forest trees may benefit from foliage removal by invertebrates. Alleviation of drought or nutrient stress through defoliation also has been shown to reduce plant vulnerability to pathogens that infect stressed trees.

Litter invertebrates also contribute to growth of forest plants. Worms and other burrowing invertebrates are critical to nutrient release from litter and to the porosity of woody litter and soil. Porous woody litter and soil facilitate root penetration and water-holding capacity.

Forest structure. Plant-eating invertebrates influence forest structure through selective herbivory on abundant or stressed hosts. Suppression of an abundant or stressed plant species opens the canopy and releases space and resources for nonhost plant species or plant life stages, often representing earlier successional stages. Because different plant species use water and nutrients differently, a diverse vegetation competes less for resources and maximizes the probability that sufficient plants will survive disturbance to maintain canopy integrity and ecosystem function.

Plant-eating invertebrates often interact with fire to facilitate nutrient turnover (especially on more arid regions), and maintain vegetation structure and diversity, often by promoting plants tolerant of drought and fire (**Fig. 1**). Fire suppression in managed forests is associated with dense, stressed trees vulnerable to outbreaks of plant-eating invertebrates. Defoliation or plant mortality stimulates turnover of plant tissues and increases fuel accumulation, leading to increased likelihood of fire, and ensures survival of the fire-adapted tree species.

Invertebrates may affect local and regional climates through their effects on forest structure. Forests with closed canopies create relatively moderate environmental conditions because of shading and interception of precipitation and wind. Defoliation or tree mortality affects climate by altering albedo (reflectance), increasing soil temperature, reducing evapotranspiration and relative humidity, increasing canopy surface roughness and wind turbulence, and stimulating emission (through decomposition) of water, carbon dioxide, and nitrous oxide (major greenhouse gases) to the atmosphere. *See* CLIMATE HISTORY; PLANT RESPIRATION.

Nutrient cycling. Invertebrates contribute substantially to nutrient cycling processes critical to soil

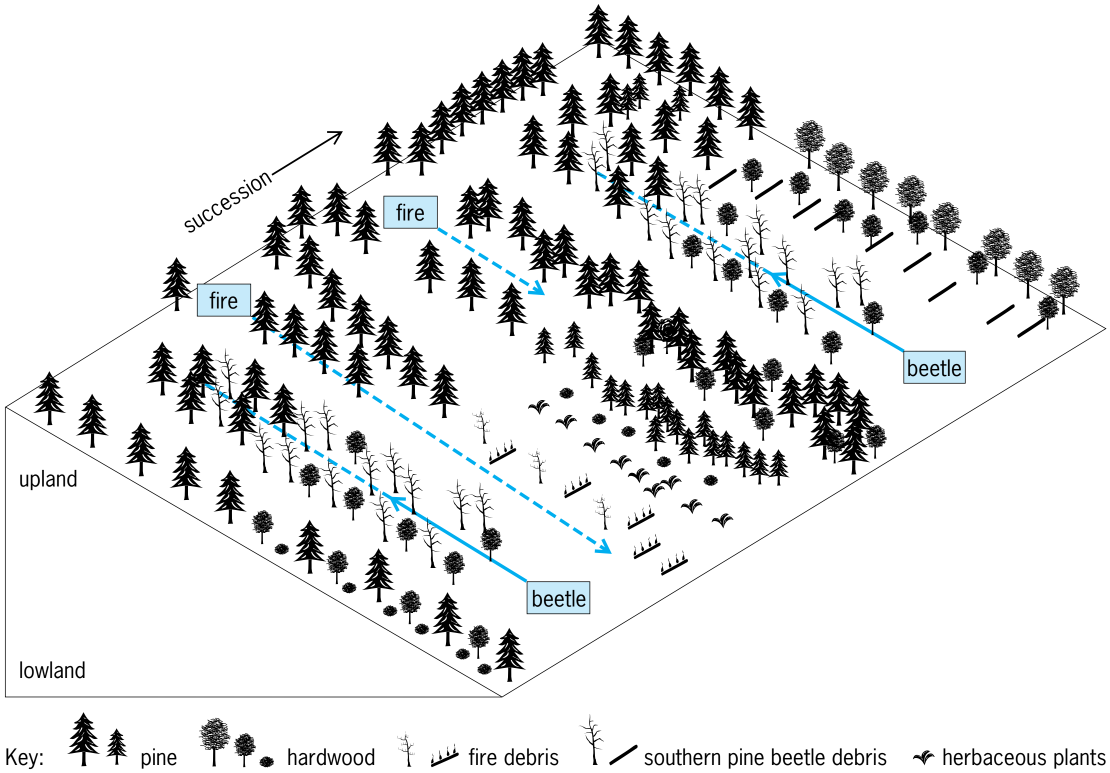


Fig. 1. Southeastern coniferous forest (upland and lowland), illustrating interactions between the southern pine beetle and fire. (*After T. D. Schowalter, R. N. Coulson and D. A. Crossley, Jr., Role of Southern pine beetle and fire in maintenance of structure and function of the southeastern coniferous forest, Environ. Entomol., 10(6):821–825, 1981*)

fertility and forest productivity (**Fig. 2**). First, herbivorous invertebrates are major regulators of nutrient turnover from plant biomass. Pruning and thinning are established horticultural methods for promoting plant growth and fruit production. Invertebrates similarly stimulate plant growth by removing plants or plant parts and thereby reducing competition for limited plant resources. Consumption and turnover of plant tissues throughout the growing season provides a more constant flow of organic matter and nutrients to the forest floor, compared to bulk litterfall prior to cold or dry seasons. More constant organic matter and nutrient supply provides resources for organisms in soil and litter and contributes to soil fertility. Partially digested plant material passed through herbivore guts is fragmented and infused with enzymes and microorganisms that facilitate chemical breakdown. Plant defensive compounds may be partially or wholly metabolized, enhancing subsequent decomposition on the forest floor. Defoliation can stimulate microbial nitrogen fixation and nitrification processes on the forest floor, processes critical to supply of this important nutrient. *See* NITROGEN CYCLE.

Invertebrates that feed an wood, especially beetles, are instrumental in initiating decomposition and nutrient turnover from dying trees and woody litter by penetrating bark and inoculating wood with saprophytic and nitrogen-fixing microorganisms. The winding galleries of wood-eating beetles and termites ensure rapid inoculation of microorganisms throughout logs. Termites host nitrogen-fixing bacteria or protozoa in their guts and increase the availability of nitrogen in the vicinity of their colonies.

Comminution (fragmentation) of dead plant or animal matter is a necessary precursor to litter decomposition and nutrient release (mineralization). A diverse array of saprophagous soil and litter invertebrates are responsible for the fragmentation and infusion of ingested food by microorganisms. Litter usually has lower nutritional value than living tissues because of the retrieval by plants of essential nutrients, especially nitrogen. Therefore, decomposition usually does not progress substantially until litter has been preconditioned by passage through invertebrate guts. Large invertebrates such as millipedes and annelid worms are responsible for fragmenting large litter material; smaller invertebrates such as oribatid mites and springtails process finer material. Most of these invertebrates also are coprophagous, reingesting fecal material after microbial enrichment.

Invertebrates also concentrate major nutrients several orders of magnitude above plant and soil-litter concentrations. Defoliators are particularly rich sources of potassium, calcium, and magnesium, and termites are rich in nitrogen. These nutrients become available to other organisms more readily than do nutrients in plant tissues.

The process of ecosystem recovery from disturbance, as affected by insects and pathogens, also contributes to nutrient balance in forest ecosystems.

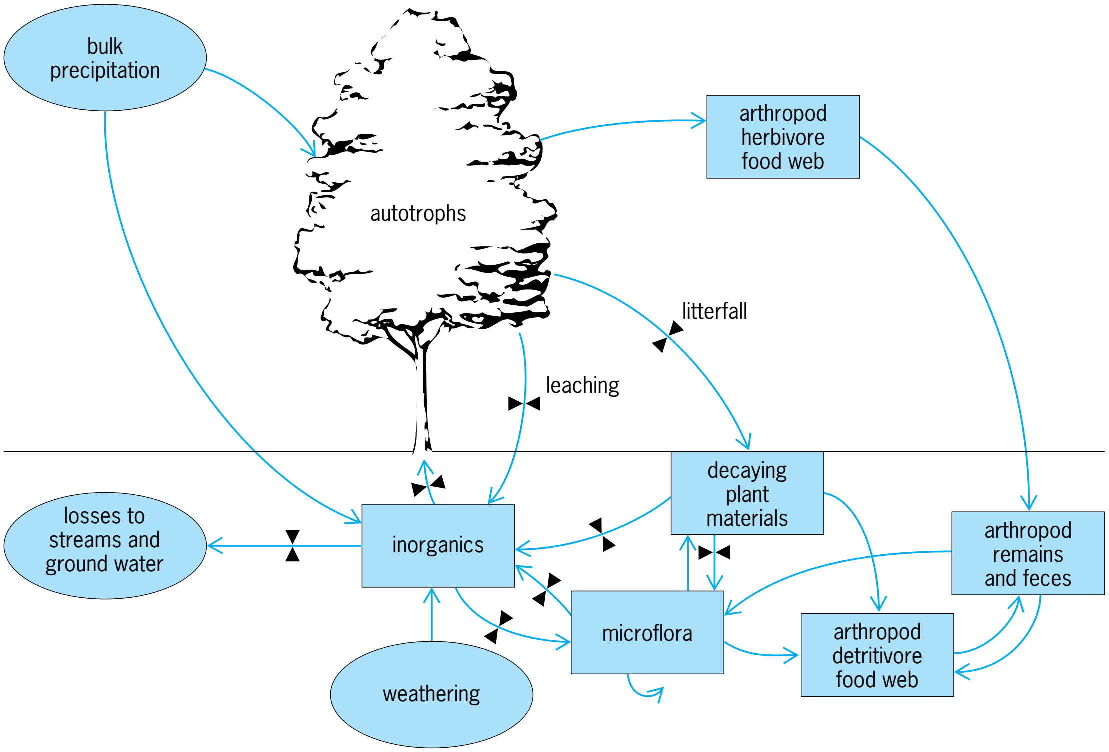


Fig. 2. Conceptual model of elemental cycling in a terrestial ecosystem, emphasizing the presence and activities of arthropod consumers. The indirect regulation of elemental movement by arthropods is indicated by the hourglass-shaped valves on these flows. (*After T. R. Seastedt and D. A. Crossley, Jr., The influence of arthropods in ecosystems, BioScience, 34(3):157–161, 1984*)

Recovery is facilitated by fast-growing early successional species that incorporate nutrients into biomass. As these species grow and later successional species become established, crowding leads to competitive stress, eventually triggering population outbreaks of plant-eating invertebrates and pathogens on the most stressed plants. Thus, invertebrates are instrumental in accelerating the successional transition to more tolerant, nutrient-conserving species, as well as opening the canopy (often over large areas) and promoting earlier successional species. Tim Schowalter

Microorganisms

Microorganisms, including bacteria, fungi, and protists, are the most numerous and the most diverse of the life forms that comprise any forest ecosystem. Their size, prolific reproduction, and varied nutritional strategies allow them to colonize most surfaces above and below ground. The structure and functioning of forests are dependent on microbial interactions. Four processes are particularly important: nitrogen fixation, decomposition and nutrient cycling, pathogenesis, and mutualistic symbiosis.

Nitrogen fixation. Nitrogen fixation is crucial to forest function. While atmospheric nitrogen is abundant, it is unavailable to trees or other plants unless fixed, that is, converted to ammonia (NH_4), by either symbiotic or free-living soil bacteria. Symbiotic nitrogen fixation takes place in root nodules on leguminous plants such as acacia, and on a few other plants including, in the forest, alder (*Alnus*) and the shrub *Ceanothus*.

Lichens are a symbiotic association between algae or cyanobacteria and fungi. Lichens with cyanobacteria fix nitrogen and provide a significant portion of the nitrogen input to old-growth coniferous forests, for example. They are very sensitive to air pollutants and are used to monitor air quality. However, their diversity and abundance decreases with proximity to sources of air pollution. *See* AIR POLLUTION; NITROGEN FIXATION.

Decomposition and nutrient cycling. Most microorganisms are saprophytic decomposers, gaining carbon from the dead remains of other plants or animals. In the process of their growth and death, they release nutrients from the forest litter, making them available once again for the growth of plants. Their roles in carbon, nitrogen, and phosphorus cycling are particularly important. Fungi are generally most important in acid soils beneath conifer forests, while bacteria are more important in soils with a higher pH. Bacteria often are the last scavengers in the food web and in turn serve as food to a host of microarthropods (**Fig. 3**).

Microorganisms reduce the mass of forest litter and, in the process, contribute significantly to the structure and fertility of soils as the organic residue is incorporated. Basidiomycete fungi are particularly important because they possess enzyme systems that allow them to obtain energy from decomposition of the complex carbohydrates (cellulose) and phenolic

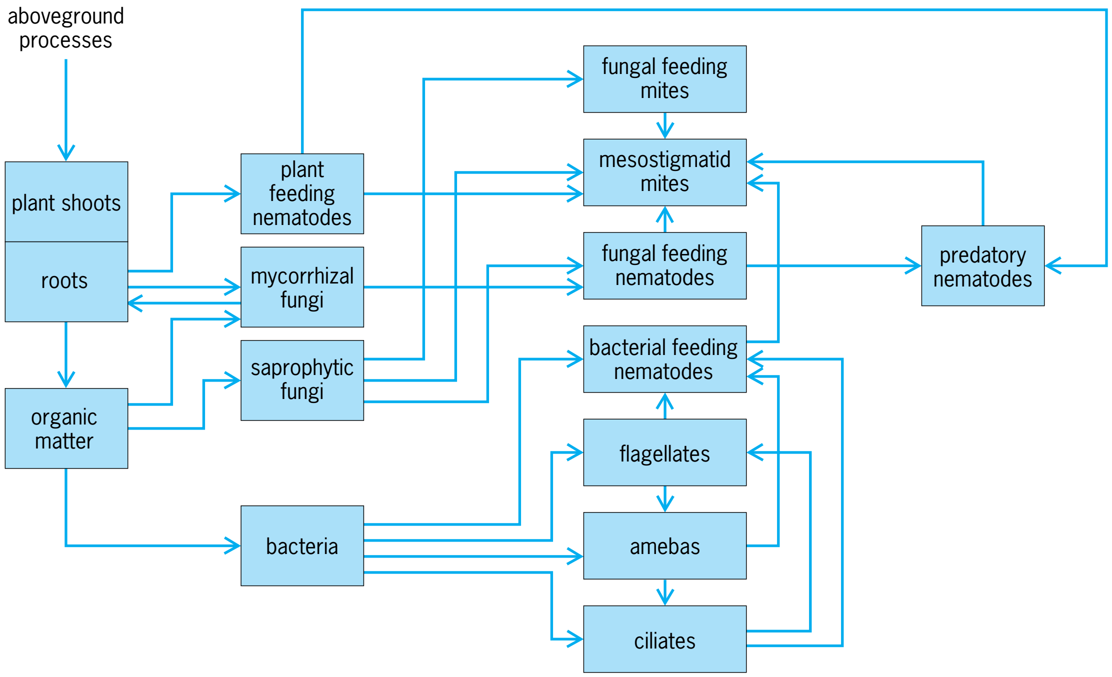


Fig. 3. Belowground food web. (*Courtesy of Elaine Ingham*)

compounds (lignin) of wood. Some fungi that can decompose lignin (white rot fungi) are used to remove or detoxify a variety of organic wastes in industrial processes.

Plant pathogens. Some bacteria and many fungi are plant pathogens, obtaining their nutrients from living plants. The species of plant-pathogenic fungi are specialized in many ways. Some are opportunists, successful as saprophytes, but capable of killing weakened or wounded plant tissues. Others require a living host, often preferring the most vigorous trees in the forest. Pathogenic fungi usually specialize on roots or stems or leaves, on one species or genus of trees. Some pathogens kill their hosts outright, but most attack only a portion of the tree, perhaps making it vulnerable to some secondary agent such as a bark beetle. Plant pathogens gain entry through wounds or by forcing or dissolving their way through the natural barriers of bark and cuticle. They secrete toxins that kill cells, allowing nutrients to escape and be absorbed, and enzymes that dissolve cell walls and allow further colonization by the fungus. Their strategies are diverse, but all benefit by colonizing living plant tissues, before the numerous saprophytes can become established.

Pathogenic fungi are important parts of all natural forest ecosystems. The forest trees evolved with the fungi, and have effective means of defense and escape, reducing the frequency of infection and slowing the rates of tissue death and tree mortality. However, trees are killed, and the composition and structure of the forest is shaped in large part by pathogens.

Pathogens remove weak or poorly adapted organisms from the forest, thus maintaining the fitness of the population. Decay fungi that kill parts of trees or rot the heartwood of living trees create an essential habitat for cavity-nesting birds and the other animals dependent on hollow trees.

By killing trees, pathogens create light gaps in the forest canopy. The size and rate of light gap formation, whether from the death of individual trees by a canker fungus or from the death of expanding groups of trees by a root rot pathogen, and the relative susceptibility of the tree species present on the site determine the ecological consequences of mortality. Succession is often advanced as shade-tolerant trees are released in small gaps. For example, in the Douglas-fir forest the late successional hemlock can reach the canopy only when the seral fir is killed, usually by root rot and the associated bark beetles. Gaps allow the growth of herbaceous plants in the island of light, creating habitat and food diversity for animals within the forest. In many forests, pathogens are the most important gap formers and the principal determinants of structure and succession in the long intervals between stand-replacing disturbances such as wildfires or hurricanes. *See* PLANT PATHOLOGY.

Mycorrhizae. The fungus roots of trees, and indeed most plants, represent an intimate physical and physiological association of particular fungi and their hosts. Mycorrhizae are the products of long coevolution between fungus and plant, resulting in mutual dependency. Two main types of mycorrhizae are known, representing distinct groups of fungi and host plants.

Ectomycorrhizae are composed of a fungus sheath around the outside of root tips, with individual hyphae penetrating between the cortical cells of the root to absorb photosynthates. Outside the root, the hyphae extend into the soil matrix, absorbing and transferring water and nutrients to the tree. In endomycorrhizae (also called vesicular-arbuscular mycorrhizae), the fungus colonizes inside the roots, forming distinctive absorptive structures within cortical cells. Again the hyphae extend into the surrounding soil.

Mycorrhizae are particularly important to trees because they enhance the uptake of phosphorus from soils. This mineral is relatively immobile in the soil and soon is depleted in the immediate vicinity of growing roots. Mycorrhizal fungi greatly extend the absorptive surface of roots through the network of external hyphae. *See* ECOSYSTEM; FOREST AND FORESTRY. Everett Hausen

Bibliography. J. R. Barnes and G. W. Minshall (eds.), *Stream Ecology*, 1983; F. H. Bormann and G. E. Likens, *Pattern and Process in a Forested Ecosystem*, 1979; G. C. Carroll and D. T. Wicklow, *The Fungal Community: Its Organization and Role in the Ecosystem*, 1992; C. C. Grier, K. J. Elliott, and D. G. McCullough, Biomass distribution and productivity of *Pinus edulis-Juniperus monosperma* woodlands of north-central Arizona, *For. Ecol. Manag.*, 50:331–350, 1992; R. T. T. Forman and M. Godron, *Landscape Ecology*, 1986; M. L. Hunter, *Wildlife, Forests, and Forestry: Principles of Managing Forests for Biological Diversity*, 1990; J. P. Kimmins, *Forest Ecology*, 1987; J. L. Meyer, A blackwater perspective on riverine ecosystems, *BioScience*, 40:643–651, 1990; C. D. Oliver and B. C. Larson, *Forest Stand Dynamics*, 1990; T. D. Schowalter, W. W. Hargrove, and D. A. Crossely, Jr., Herbivory in forested ecosystems, *Annu. Rev. Entomol.*, 31:177–196, 1986; J. A. Standford and J. V. Ward, The hyporheic habitat of river ecosystems, *Nature*, 335:64–66, 1985; E. O. Wilson, *The Diversity of Life*, 1992.

Forest fire

The term wildfire refers to all uncontrolled fires that burn surface vegetation (grass, weeds, grainfields, brush, chaparral, tundra, and forest and woodland); often these fires also involve structures. In addition to the wildfires, several million acres of forest land are intentionally burned each year under controlled conditions to accomplish some silvicultural or other land-use objective or for hazard reduction (**Fig. 1**).

Fig. 1. Controlled fire conditions. (*a*) Closed jack pine (*Pinus banksiana*) cones, which require heat in order to open. (*b*) Prescribed burning after seed tree cutting so that jack pine cones can open for natural regeneration. (*Photographs by W. R. Beaufait*)

Nature of fire. Fire is a natural phenomenon that reverses the creative process of photosynthesis, which produces all wood fiber and other plant materials. In simple terms, photosynthesis uses solar energy to turn carbon dioxide and water into cellulose (woody fiber) and oxygen: $CO_2 + H_2O + \text{solar energy} \rightarrow \text{cellulose} + O$. When the temperature near woody or other plant material is raised to the kindling point by a lighted match, burning cigarette, or other means, combustion takes place. As the material burns, carbon dioxide and water are released, along with heat which can spread the fire to nearby combustible plant materials. If dead woody or plant material does not get burned, it decays. In decay the same process produces the same end products of carbon dioxide, water, and heat at a much slower pace. The heat of oxidation is usually not noticeable, and the debris of decaying organic material accumulates to form a thick layer, as on the forest floor. During dry seasons this layer may dry out to such an extent as to be a serious potential fire hazard. *See* COMBUSTION.

Forest fire causes. Most wildfires are caused by human beings, directly or indirectly. In the United States less than 10% of all such fires are caused by lightning, the only truly natural cause. In the West (the 17 Pacific and Rocky Mountain states) lightning is the primary cause, with smoking (cigarettes, matches, and such) the second most frequent. Combined they account for 50 to 75% of all wildfires. In the 13 southern states (Virginia to Texas) the primary cause is incendiary. This combined with smoking and debris burning make up 75% of all wildfires. Miscellaneous causes of wildfires are next in importance in most regions. The other causes of wildfires are machine use and campfires. Machine use includes railroads, logging, sawmills, and other operations using equipment.

Fire behavior. The manner in which fuel ignites, flame develops, and fire spreads and exhibits other phenomena constitutes the field of fire behavior. Factors determining forest fire behavior may be considered under four headings: attributes of the fuel, the atmosphere, topography, and ignition.

Fuel. Live and dead vegetation in the natural forest complex constitutes the fuel for forest fires. The amount of moisture in relation to the heat content of the fuel determines its flammability. Moisture in living vegetation varies with species and with time of year. The moisture content of dead fuels, the more important component of the fuel complex, responds to precipitation, relative humidity, and to a lesser extent temperature. It shows important daily fluctuations. *See* WOOD PROPERTIES.

The rate of combustion and the heat output are also determined by the quantity of fuel, the fuel arrangement, and the thermal and chemical properties of individual fuel particles.

Atmosphere. Major direct effects of the atmosphere on fire behavior are through variations in oxygen supply and in flame angle caused by wind, and the stability of the atmosphere. Secondary effects on

Fig. 2. Characteristic fire behavior. (*a*) Crown fire. (*b*) Surface fire. (*U.S. Forest Service*)

combustion rate are through air and fuel temperature, and the relative humidity of the air. *See* ATMOSPHERE; HUMIDITY.

Topography. Heated air adjacent to a slope tends to flow up the slope, creating a chimney effect. Unburned fuels above a fire advancing up a slope are closer to approaching flames because of the angle of the slope. Both of these effects act similarly to wind on flat ground in increasing the rate of fire spread.

Ignition. An igniting agent of sufficient intensity and duration is necessary to start a forest fire. The agent may be human-caused, such as a match or lighted cigarette, or natural, such as lightning. Once a fire has started, its spread is by successive ignition of adjacent unburned fuel. This fuel is brought to ignition temperature by heat radiated and convected from the flame front. Generally fuels must be very close to, or even enveloped by, flame or superheated gases in the convection column before igniting. This mechanism results in more or less continuous advancement of a fire. Fire spread may also be caused by spotting, that is, ignition by burning brands carried outside the fire area by wind and turbulence.

Types. A forest fire may burn primarily in the crowns of trees and shrubs—a crown fire (**Fig. 2***a*); primarily in the surface litter and loose debris of the forest floor and small vegetation—a surface fire (Fig. 2*b*); or in the organic material beneath the surface litter—a ground fire. The most common type is a surface fire.

Indexes. The U.S. Forest Service has developed a National Fire Danger Rating System (NFDRS) to provide fire-control personnel with numerical ratings to help them with the tasks of fire-control planning and the suppression of specific fires. The system includes three basic indexes: an occurrence index, a burning index, and a fire load index. Each of these is related to a specific part of the fire-control job. These indexes are used by dispatchers in making decisions on setting up fire-fighting forces, lookout systems, and so forth.

Seasons. Because of annual cycles in precipitation and in growth of vegetation, most sections of the United States have a well-defined fire season during which most forest and wild-land fires occur. This fire season varies widely in duration, from two short seasons in the northeastern states, in spring and fall, to frequently a year-long season in southern California. **Figure 3** shows the normal peak fire seasons for the United States.

Prevention of wildfire. This means keeping all unwanted human-caused wildfires from starting. Actually prevention covers a wide range of activities, and must be patterned to the causes of fires in a particular area. The best tools for prevention are education, elimination of hazards, and fire law enforcement. Risks refer to the presence of people and to their activities, such as cigarette smoking and starting brush fires. Hazards are the various types, conditions, volumes, and locations of fuels, such as grass, brush, and forest fuels.

In areas of intensive use, hazard reduction includes many forestry activities aimed at reducing fuel accumulations in order to reduce the channel of a starting fire or to reduce the severity of one that might be started. These activities include roadside burning, slash disposal, controlled burning of forest areas to reduce fuel accumulations, and some silvicultural practices. *See* SILVICULTURE.

Prevention activities should be conducted in line with a carefully planned and written program. The

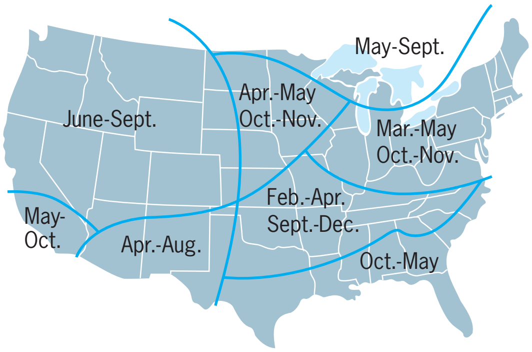


Fig. 3. Normal peak fire seasons in United States.

keystone of such a program is a wildfire prevention plan that includes at least the following elements: (1) basis of the plan (consisting of maps and statistics covering past fires, maps showing fire risks, industrial operations, and hazards, and maps showing signs, radio and television stations, and newspaper offices); (2) summary of objectives, problems, and action to be taken; (3) prevention contact plan; (4) public education; (5) closures and restrictions; (6) industrial operations; (7) reduction of physical hazards; (8) sign-posting plan; (9) prevention training; and (10) legislative needs and law enforcement.

Detection. Although many forest fires are detected and reported by local residents and transient forest users, primary reliance for detection is placed on specialized detection systems. Most fire-control agencies use a combination of lookout towers with trained observers during the fire season and regularly scheduled aerial patrols. Use of aircraft for routine detection patrol has increased greatly, paralleling the increase in aircraft use in other protection activities.

Infrared scanners developed originally by the military for battlefield applications have been adapted for use in detecting forest fires. Aircraft-mounted detectors are routinely used to detect fires started by intense lightning storms. Such scanners have the capability to detect even campfires from an altitude of 12,000 ft (3.7 km). Other electronic equipment, including television, radar, and sferics (to detect lightning discharges and thus storms), has been used experimentally.

Communication. Rapid and accurate communication is essential not only to the detection network but also to the forces engaged in actual suppression work. Land-line telephones are being rapidly supplanted or supplemented by radios. In many organizations all motorized units are equipped with two-way radios. As the size of portable, self-contained equipment is reduced, more and more supervisory personnel on actual fires are also being equipped.

Presuppression. In spite of careful plans and efforts to prevent wildfires, they do occur. Then they must be detected and reported promptly to the suppression forces responsible. These will be prepared to act if adequate suppression steps have been taken. Aside from prevention, the steps include all the preparation, organization, development, and maintenance of equipment, as well as planning, cooperation, mutual aid arrangements, and training.

Generally, public agencies are responsible for presuppression planning and suppression action. State and federal forestry or forest fire organizations have a clearcut understanding and lines of authority as to which areas are their primary responsibility. Each should have a well-defined organizational structure and a specific plan to prevent fires and to detect and suppress any fires that do occur. This plan includes maps of the type described for a fire-prevention program, an inventory of equipment, and a program of training and a program for securing additional personnel equipment if needed. The program for additional personnel should be established through the use of cooperative agreements well before each fire season begins.

Transportation. Emphasis on fast attack of discovered fires has steadily increased over the years. Many fire-control organizations, both urban and forest, divide their protection area into zones of allowable attack times. These are based on times shown by experience as necessary for successful control according to acceptable standards. Location of suppression forces and methods of transportation are planned to meet these objectives.

Because of the great distances involved, forest fire-control organizations have placed special emphasis on increasing the speed and mobility of their forces. Most personnel and equipment arrive at the majority of fires by a combination of ground-transportation methods, including the most elementary of all, walking. Therefore in areas of high value and high fire incidence, an extensive network of roads and trails is usually maintained. The development of reliable, lightweight gasoline engines has made possible the use of trail scooters and motorized equipment carriers. These require trails intermediate in quality between horse and foot trails and the more expensive truck trails.

In remote areas with relatively low and widely scattered fire incidence, the airplane and helicopter are the primary means of delivering both personnel and material to the scene of a fire. Even in areas with good road and trail networks, it is often cheaper and faster to supply suppression forces by air. Successful parachute jumping by fire fighters, started by the U.S. Forest Service in 1940, demonstrated the feasibility of wartime use of paratroops. Postwar development by the Forest Service of techniques and equipment for delivery of personnel and material by parachute and by helicopters led to increased use of airborne fire fighters in the forests of the western United States (**Figs. 4** and **5**).

Mapping. Specialized heat-sensitive infrared mapping devices are used operationally to map fire perimeters from the air. These are somewhat similar to television systems that are sensitive to radiation in the far infrared rather than to visible light. They have the great advantage of being able to see through the smoke of a fire and locate the regions of active combustion on the ground. The image, which can be photographed, indicates the temperature of the fire in shades from black to white and thus enables fire-control personnel to map the fire edge, determine the intensity of the fire along the perimeter, and locate spot fires outside the main fire at night or when the fire itself is completely obscured by smoke.

Such sophisticated systems of mapping are likely to be applied primarily in the West, where fires may involve extensive areas or the most valuable stands of virgin timber or second growth. For most ground fires that are common in the eastern states, fire maps can be quickly plotted directly on prints of aerial photographs, or made from compass and pacing surveys.

Suppression methods. Fire can be suppressed by robbing it of its fuel or oxygen supply, or enough heat

Fig. 4. Fire fighters rappelling to the scene of a fire.

may be removed to stop the chain reaction, and the fire goes out. Combustion may also be suppressed by the addition of small amounts of certain chemicals to the combustion zone.

Fuel removal. One of the most common methods of controlling a forest fire is by creating a fuelless barrier. This may be done by digging a trench to mineral soil ahead of an approaching fire of sufficient width to prevent the fire from crossing (**Fig. 6***a*). Since this may require a line of considerable width (some evidence is available that this width is approximately equal to the square of the height of the flames), it is customary for the unburned fuel between the prepared line and the approaching fire to be intentionally fired, thereby effectively widening the line.

Techniques of constructing fire lines vary with terrain, soil character, and type of vegetation. Organized crews with hand tools, such as axes, shovels, rakes, and other specially constructed tools, are used in many areas, especially those too rugged or remote to permit use of mechanized equipment (Fig. 6*b*). Specially developed motor-driven flails, similar to garden-tractor rotary tillers, are sometimes employed, although their usefulness is limited to terrain and soil types on which they can be maneuvered satisfactorily. Bulldozers are used extensively (Fig. 6*c*), and tractor-drawn plows are standard line-construction equipment in the relatively flat terrain of the southeastern states.

Occasionally, existing roads or previously prepared firebreaks are used as control lines, usually as a line from which a counterfire is set. Natural firebreaks, such as rocky ridges or rivers, are used similarly.

Heat removal. The second major method used in suppressing fires is to remove enough heat to break the chain reaction. Water, with its high specific heat and heat of vaporization, is an ideal substance for this purpose. Long the favorite with urban fire-control organizations, water is becoming more popular with forestry organizations as techniques and equipment for handling it are improved. An extraordinarily large amount of burning material can be extinguished with a small amount of water when it is used with maximum efficiency.

Hand-operated pumps with a 5-gal (14-liter) tank carried on the back are most satisfactory when used on fires in light fuels, for suppressing small spot fires, and for mop-up, the tedious work of extinguishing burning fuel after the fire has been contained within a fireproof boundary.

Gasoline-operated pumps are used extensively wherever water supplies are convenient or road networks permit the use of tank trucks (Fig. 6*d*). Portable pumps are available that can be carried by one person (Fig. 6*e*), making hose lays of several thousand feet (1 ft = 0.3 m) or even several miles (1 mi = 1.6 km) feasible. Techniques have been developed for laying hose rapidly from helicopters,

Fig. 5. Parachutes used for the delivery of fire fighters and materials to fire areas. (*a*) Two fire fighters descending. (*b*) Air drop of fire-fighting equipment, showing drop location marker (X). (*U.S. Forest Service*)

Fig. 6. Land fire suppression methods. (*a*) Digging a trench ahead of an approaching fire. (*b*) Constructing a fire line to stop the advancing fire. (*c*) Bulldozer making fire lane near the edge of the fire. (*d*) Using a tank truck to suppress fire. (*e*) Fighting fire with hose attached to a portable pump. (*U.S. Forest Service*)

cutting the time required to lay a mile of hose from hours to minutes, even seconds. For use in inaccessible regions, small units including a pump, a 50- or 100-gal (179- or 379-liter) tank, and several hundred feet of hose have been developed that can be delivered by helicopter to the scene of a fire.

Chemicals that inhibit the combustion process are used regularly in the suppression of large forest fires. These are generally prepared in the form of water solutions or slurries, and are dropped from aircraft or applied to fires from ground-based tank trucks. More than 15,000,000 gal (57×10^6 liters) of such fire-retardant chemicals are used annually in the United States.

Although surplus World War II aircraft were originally used in the development of airborne delivery systems (**Fig. 7***a*), newer aircraft have permitted more flexible and efficient release systems. Four-engine cargo planes with pressurized tanks permit the delivery of as much as 3000 gal (11,000 liters) of retardant mixture to a fire. The retardant is released at a low altitude, either directly on the burning material at or near the fire edge, or some distance in front of an advancing fire. A single pass may cover an area several yards wide and several hundred yards long. Helicopters are also used, especially in initial attack and mop-up operations (Fig. 7*b*). While helicopters carry smaller amounts of retardant materials (50 to 100 gal or 179 to 379 liters), their great accuracy makes them excellent for strategic use of expensive materials.

Retardant drops are used primarily to cool down hot spots and help hold a fire pending arrival of ground forces, and to control spot fires outside the main fire perimeter. The air tanker has limited value in attacking the head of a rapidly moving crown fire.

Little direct use is made of the third possibility of removing the oxygen supply to a fire, although water has this effect to some extent and soil is sometimes used to smother flames.

Suppression strategy. The manner in which the various suppression techniques are utilized on a fire is determined by the fire boss, who is comparable to the fire chief in urban organizations. It is the responsibility of the fire boss to determine the strategy appropriate to the particular fire. Depending on the size of the fire, the fire boss may have a few people or a few hundred, and occasionally more than a

Fig. 7. **Aerial fire suppression. (*a*) Mixing and loading fire retardants in preparation for an air drop. (*b*) A helicopter with a detachable 40-gal (151-liter) tank drops borate slurry on a pinpoint target. (*U.S. Forest Service*)**

thousand, to assist in carrying out the suppression strategy.

When a wildfire has been suppressed, it is essential that it be covered by a report. The fire boss should have someone specifically charged with report duties, or make the report personally. Usually there are printed forms available which include at least the following: (1) name of the fire or district number in which the fire occurred, (2) location of the fire, (3) size of the fire, (4) cause of the fire, (5) estimate of damage in dollars, and (6) date and times when the fire started and when it was suppressed or brought under control. *See* FOREST AND FORESTRY.

Willard S. Bromley

Bibliography. A. A. Brown and K. P. Davis, *Forest Fire: Control and Use*, 1973; T. T. Kozlowski and C. E. Ahlgren, *Fire and Ecosystems*, 1974; W. E. Reifsnyder, *Proceedings of the 10th Annual Tall Timbers Fire Ecology Conference*, 1970; T. Van Nao, *Forest Fire Prevention and Control*, 1982.

Forest genetics

The subdiscipline of genetics concerned with genetic variation and inheritance in forest trees. The study of forest genetics is important because of the unique biological nature of forest trees (large, long-lived plants covering 30% of the Earth's land surface) and because of the social and economic importance of forests. Knowledge of forst genetics is also important. Forest genetics is the basis for conservation, maintenance, and management of healthy forest ecosystems; and development of programs which breed high-yielding varieties of commercially important tree species.

Variation in natural forests. The outward appearance of a tree is called its phenotype. The phenotype is any characteristic of the tree that can be measured or observed such as its height or leaf color. The phenotype is influenced by (1) the tree's genetic potential (its genotype) determined by genes residing in every living cell in the tree; and (2) the environment in which the tree grows as determined by climate, soil, diseases, pests, and competition with other plants. Thus, a tree may be taller than its neighbor because it has the genetic potential for faster growth or because it is growing in a slightly better environment with more light or better soil (or both).

No two trees of the same species have exactly the same phenotype, and in most forests there is tremendous phenotypic variation among trees of the same species (for example, trees of different size, shape, branching patterns, and leaf color; **Fig. 1**). It is only on a very homogeneous site planted with trees of the exact same genotype (called a clone) that the trees may appear to have little phenotypic variability, and even then phenotypic differences are found. Forest geneticists often question whether the observed phenotypic variation among trees in forests is caused mostly by genetic differences or by differences in environmental effects. In other words, which is more important, nature or nurture?

It is impossible to look at the outward appearance of a tree (its phenotype) and know anything about its underlying genotype, because the genetic and environmental effects are combined in the phenotype. For example, a tree can be tall because of its genetic potential or its environment or both. Therefore, common garden tests are often used to hold the environment constant and therefore isolate the

***Eucalyptus grandis* is shown growing in two different conditions that exhibit different levels of phenotypic variability. (*a*) In a natural stand in Australia with large phenotypic variability caused by both genetic and environmental differences among the trees (*photo by K. Eldridge*). (*b*) In a plantation of Mondi Forests in South Africa where the trees are more uniform both because the site is uniform and because all trees are the same genotype or same clone (*photo by T. White*).**

genetic and environmental effects on phenotypic variability. To accomplish this, seed is collected from many trees of the same species growing in natural forests, and these seed are planted in randomized, replicated experiments in one or more common field locations (called common gardens). Under these conditions, if the seedling offspring (called progeny) from some parent trees consistently perform better than progeny from other parents, these average differences must be due to genetic causes since the progeny from all parents are being compared in the same common environments.

As an example, suppose that the heights of 1000 trees are measured in a forest of pure loblolly pine, *Pinus taeda*. Each tree has its own genotype, and each has experienced its own microclimate and microsite in the unique part of the stand where it grows. Thus, the measurements of 1000 phenotypes produce a wide range of tree heights, from short to tall. This is the phenotypic variation for the particular trait of height growth. A common-garden experiment to assess how much of this phenotypic variation is caused by genotypic differences among the 1000 trees might be established as follows: (1) Collect seed from the tallest five trees and the shortest five trees in the stand. (2) Keep the seed from each tree in separate bags so that there are ten bags or ten treatments. (3) Plant the seedlings produced from the ten different treatments in randomized, replicated designs in several forest locations (these new field sites are the common gardens). (4) After several years, measure the heights of all seedlings in all replications on all sites. (5) Condense the data by calculating ten averages, one for each of the ten treatments corresponding to the ten parent trees in the original stand. If the progeny seedlings originating from the five tall trees in the parent stand are taller, on average, than the progeny from the five short parents, then at least a portion of the original phenotypic variability among the ten parents must have been caused by genetic differences. This genetic portion of the phenotypic superiority for height growth of the taller parents was passed on to the offspring, as proven by the superior growth of their progeny in common-garden environments.

Geographic variation in natural forests. The term "provenance" refers to a specific geographical location within the natural range of a tree species. Genetic differences among provenances can often be quite large, especially for wide-ranging species occupying many diverse climates. For example, Norway spruce (*Picea abies*) has an extremely wide natural distribution in Europe and Asia that spans many different countries, elevations, climates, and soil types. Natural selection during the course of evolution has adapted each provenance to its particular local environment. This means that there are large genetic differences among provenances growing in different environments. Provenances originating from colder regions (farther north or higher elevations) tend to grow more slowly, begin height growth earlier each spring, end height growth earlier each fall, and have narrower crowns with flatter branches better adapted to the dry snow and types of frosts in colder climates. To demonstrate that these differences among provenances are genetic in origin, common-garden tests, called provenance tests, have been planted. That is, seed has been collected from several provenances and planted for comparison in randomized, replicated studies in various forest locations.

Most wide-ranging species of forest trees (such as Douglas-fir, *Pseudotsuga menziesii*, in the western United States, and loblolly pine, *Pinus taeda*, in the eastern United States) have large genetic differences among provenances from different parts of the range. However, common-garden experiments have revealed that natural selection also acts on a smaller scale within any given region. For example, Douglas-fir stands at lower elevations in the Cascade Mountains of Oregon are genetically distinct from higher-elevation provenances only 50 mi (80 km) apart. Also, Douglas-fir trees from the south and north side of the same mountain can sometimes differ genetically, presumably in response to natural selection on the colder north-facing slopes.

Some tree species show relatively little geographic variation among provenances. Sometimes this is because the native range is relatively uniform in its climate and soils, such that there have not been strong differences in natural selection pressures in different parts of the range (such as slash pine, *Pinus elliottii*, in the southeastern United States).

The study of geographic variation through provenance tests should be a first step in the genetic research of any tree species. The information gained about the amount and patterns of genetic variation among provenances within the species' native range is important for (1) learning about the interplay and importance of evolutionary forces that caused the observed patterns; (2) making reforestation decisions about how far seed can safely be transferred from the site of seed collection to a distant planting site and still ensure adequate adaptation; (3) determining the most suitable provenances to use to initiate a tree breeding program for the species; and (4) formulating gene conservation strategies that capture the natural genetic diversity that exists within the species.

Genetic variation within forest stands. In addition to genetic differences among provenances, there is usually substantial genetic variation among trees within the same provenance and even within the same forest stand. There are two reasons for this genetic diversity: (1) Different trees have different genotypes in most natural stands. (2) Each tree is heterozygous for many genes, meaning that a given tree has multiple forms (called different alleles) of many genes. Population genetics studies patterns of genetic diversity in populations (such as forest stands). Results of many studies have shown that most forest tree species maintain very high levels of genetic diversity within populations.

Most tree species do not respond well to self-pollination or other forms of inbreeding which would lead to a reduction in genetic diversity. Rather, most species of forest trees have evolved one or more mechanisms to promote outcrossing (that is, cross-pollination between different trees) and maintain genetic diversity: (1) In many conifers (such as pines), the female flowers (that subsequently mature into pine cones) are located in the tree tops, and the pollen-producing catkins (male parts) are located farther down in the tree crown, and this spatial separation helps avoid self-pollination. (2) In some tree species, the male and female parts are ripe for pollination at different times, and this temporal separation of several days helps avoid self-pollination. (3) Even after self-pollination, there are sometimes molecular and physiological mechanisms that favor the germination, growth, and survival of outcross pollen, thus tending to inhibit self-fertilization.

While most tree species maintain large amounts of genetic diversity, there are some exceptions. One example is Torrey pine, *Pinus torreyana*, from southern California, which shows almost no genetic tree-to-tree variation; that is, all trees are nearly the same genotype. In the evolutionary past of these species with little genetic diversity, there have been severe reductions in the total number of living trees of the species (such as due to advance of glaciers during the ice ages). These severe reductions in population size eliminate much of the genetic diversity of the species. Even if the native range of the species today is substantial with many millions of trees, all trees are descendants of the few trees left after the reduction; hence, the genetic diversity is low. *See* POPULATION GENETICS.

Forest tree breeding. Beginning with the natural genetic variation that exists in an undomesticated tree species, tree breeding programs use selection, breeding, and other techniques to change gene frequencies for a few key traits of the chosen species. As with agricultural crops, tree breeders produce genetically improved, commercial varieties that are healthier, grow faster, and yield better wood products. After an existing forest stand is harvested, a new stand of trees is planted to replace the previous stand in the process called reforestation. Use of a genetically improved variety for this reforestation means that the new plantation will grow faster and produce wood products sooner than did the previous stand.

Large-scale tree breeding programs began in earnest in the 1950s. Today, tree breeding programs are very widespread in the world. There are breeding programs for all forest tree species that are planted in substantial quantity. Depending on the species and country, some programs are conducted by single private companies, others by government agencies, and still others by many organizations in which several companies and agencies pool their efforts. In the southeastern United States alone, there are more than 30 private companies and state forestry agencies working cooperatively to breed improved varieties of loblolly and slash pine. Each year these organizations produce more than 1 billion genetically improved seedlings that are used to reforest more than 1 million harvested acres (400,000 hectares). This is more than 90% of the total annual reforestation with pines in the southeast.

Due to the long time frames involved and their recent initiation, tree breeding programs have not advanced as far as breeding programs for most agricultural crops. This means that the new improved tree varieties are usually not discernibly different to the naked eye from the undomesticated, natural species. Only a few traits of key importance (such as growth rate, disease resistance, and wood properties) are targeted for genetic improvement in the tree breeding program, and there is little change in most other traits. Nevertheless, these new varieties benefit the organizations involved by producing more wood products in a shorter time. On a global scale, the overall benefit to society of tree breeding programs is to increase the yields from the world's forest plantations. This helps to meet the rising demand for forest products, which in turn reduces the need to meet this demand from natural forests. Thus, more natural forests can be conserved or used for other social objectives.

Forest biotechnology. Several laboratory techniques of molecular genetics and biotechnology promise to make major contributions to forest genetics and tree breeding in the near future: (1) Somatic embryogenesis is a technique to duplicate (or propagate) selected trees asexually from their vegetative (somatic) cells, and this allows the best trees to be immediately propagated commercially as clones without sexual reproduction (that is, no seed is involved). (2) Genetic mapping of some important tree species is well under way, and these maps will be useful in many ways to understand the genetic control of important traits, such as disease resistance. (3) Marker-assisted selection is the use of some kinds of genetic maps to help select excellent trees at very early ages based on their deoxyribonucleic acid (DNA) genotype as assessed in a laboratory (instead of growing trees in the field and selecting based on performance in the forest). (4) Functional genomic analysis is an exciting new field of genetics that aims to understand the function, controlled expression, and interaction of genes in complex traits such as tree growth. (5) Genetic engineering or genetic modification is the insertion of new genes into trees from other species, and as one example the gene for resistance to the herbicide glyphosate has been inserted into species of *Populus* and *Eucalyptus*.

Conservation of genetic resources. For commercially important species of forest trees, gene conservation is practiced by tree breeding programs to sustain the genetic diversity needed by the program. However, conservation of genetic diversity is a major global concern and is important for all forest species to (1) maintain the health and function of forest ecosystems, and (2) sustain the genetic diversity of noncommercial species that may eventually have economic value. There are many current medicinal and other nonwood products produced by trees (especially trees in tropical forests such as in the Amazon), and there are many more future products as yet undiscovered. For example, the Pacific yew (*Taxus brevifolia*) native to the Pacific Northwest of the United States was once considered a weed. However, this species recently became valuable for its production of a natural chemical, taxol, used in fighting certain types of cancer.

There are two broad categories of gene conservation programs. In situ programs conserve entire forest ecosystems in forest reserves, national parks, wilderness areas, or other areas set aside for conservation purposes. Ex situ programs obtain a sample of the genotypes from different provenances of a single tree species and collect seed or vegetative plant material from each genotype to store in a separate location (such as a seed bank in a refrigerated room). Both types of programs are important for conserving the world's forest resources. *See* BREEDING (PLANT); FOREST AND FORESTRY; FOREST ECOSYSTEM; GENETIC ENGINEERING; PLANT PROPAGATION. Timothy L. White

Bibliography. A. K. Mandal and G. L. Gibson (eds.), *Forest Genetics and Tree Improvement*, CBS Publishers, New Delhi, 1998; T. L. White, W. T. Adams, and D. B. Neale, *Forest Genetics*, CAB Publishers, London, 2001; B. J. Zobel and J. Talbert, *Applied Forest Tree Improvement*, John Wiley, New York, 1984.

Forest harvest and engineering

Application of engineering principles to the solution of forestry problems, such as those dealing with harvesting, forest transportation, materials handling, and mechanical silviculture, with regard to long-range environmental and economic effects.

The work that forest engineers perform varies widely throughout the United States. In the Northeast and Southeast, tasks include mechanization of harvesting, site preparation, planting, and product handling. Development and modification of machinery is an important part of this job, as is the improvement of the ability of workers to use machines in the woods, an activity known as work science. In the West, the terrain changes the job, as does the size of the trees usually harvested. Planning, design, and construction of road systems are major operational and environmental challenges in the West. Harvesting with mechanical fellers or yarding with ground skidders is often limited by tree size and terrain. Cable systems are commonly used to overcome these constraints, and the design and layout of these cable logging units require a great deal of engineering skill if logging is to be done in an economic and environmentally acceptable manner. In short, the skills possessed by forest engineers in the eastern United States parallel most closely those of the mechanical or agricultural engineer. In the West, the skills are more closely aligned with civil engineering. *See* AGRICULTURAL ENGINEERING; CIVIL ENGINEERING; MECHANICAL ENGINEERING.

Education. The education of the forest engineer in the United States and in most of the universities teaching this subject matter throughout the world is most complex. The profession is comparatively new; it grew out of the need for professional engineers in forestry. It has been described as a mixture of the traditional forest management and civil engineering curricula. The forest engineering curriculum had its beginning along these lines, but since the early 1950s, at least in the United States, it has developed its own curriculum, including a number of courses that are oriented toward the objectives of the forest engineering profession. Obviously, curricula vary from school to school, but a representative university curriculum would include courses on general, agricultural, and civil engineering; forest engineering; forest management and forest products; physical, biological, and mathematical sciences; social sciences; economics; business; and the humanities and the arts.

Most curricula are based on four years of study for the bachelor's degree, with options for a five-year double degree in forest engineering–forest management, forest engineering–civil engineering, and so on. Students can proceed to the master's degree and doctor's degree in forest engineering. Summer field work or a summer camp is required at most schools of forestry.

In several states, forest engineering graduates are eligible to take the state land surveying-in-training and the engineer-in-training tests upon graduation. If they pass, they are eligible to take the state logging test, and some qualify for the civil engineer's test as well.

Since forest engineering is very specialized, most of the graduates elect to stay within the field. The majority find private employment, while the remainder find positions in the U.S. Forest Service, Bureau of Land Management, and the various state and county forestry services.

Logging practices. Forest engineering encompasses many important and diverse subjects. Logging is one of the most important aspects. It is very critical that the logging engineer not only log at the lowest cost possible but also consider other items, such as safety, water and soil protection, ergonomics, esthetics, machine design and development, and systems development.

Generally, the definition of logging includes the planning stage, felling and bucking, yarding, loading, hauling, and unloading. The most spectacular of these and usually the most expensive is the yarding or skidding phase. Yarding is moving a log by some method, usually a cable system, that applies some lift to the leading end of the log; skidding accomplishes the same result, but very little or no lift is applied. A typical example of skidding would be horse skidding, during which the log is in contact with the ground at all times. Yarding or skidding methods fall into the following categories: animal skidding, cable yarding, surface-vehicle yarding or skidding, and under exceptional conditions helicopter or balloon yarding. *See* LOGGING.

Forest engineering problems. The specialized training and subsequent experience of qualified forest engineers are called upon when they may be required to do the following:

1. Compare the costs of owning and operating different units and different combinations of logging equipment. Such comparisons are usually made by dividing cost per unit of time by the production rate— the volume of board-feet, cords, or other volume produced for the same unit of time.

2. Recommend the standards of the roads that are to be used in transporting forest products from the woods.

3. Determine the most efficient spacing of logging and other roads with a given set of logging equipment, or with a given plan for managing all the forest resources.

4. Survey and designate the location of logging and other roads.

5. Locate and design the type of culverts, bridges, and other construction features of the road system recommended.

6. Assist with the location, design, and construction of recreational trails, buildings, shelters, camps, and picnic and playground facilities.

7. Assist with the selection and operation of the most efficient machinery to be used in such silvicultural programs as forest tree nursery operations, site preparation, tree planting, and management of problem vegetation.

8. Be responsible in all planning and operation to conform with environmental regulations and all laws affecting the equipment and programs under the forest engineer's recommendations or direction.

9. Be conscious of and provide for the safety of the persons operating the equipment or using the structures recommended or designed.

Environmental aspects. Protecting the forest environment is increasingly important to the forest engineer. The "cut and get out" attitude toward forest properties is a thing of the past. Today's engineer has an obligation to the employer as well as the public to preserve forests, whether they are publicly or privately owned. Sometimes, this creates a conflict among interested parties in various phases of forest management. Environmental issues include water quality, soil disturbance, preservation of scenic values and recreation and wilderness areas, presence of wildlife, controversial logging methods like clearcutting, smoke due to slash burning, and forest pest control. *See* FOREST AND FORESTRY; FOREST MANAGEMENT.

John E. O'Leary

Bibliography. F. C. Ford-Robertson and R. K. Winters (eds.), *Terminology of Forest Science, Technology, Practice and Products,* rev. ed., 1983; B. H. Maddison, *Handbook of Timber Engineering Design,* 1982; K. A. Staaf (ed.), *Tree Harvesting Techniques,* 1984.

Forest management

The planning and implementation of sustainable production of forest crops and other forest resources and uses. Key decisions include land allocation to different uses or combination of uses, silvicultural method and practices, intensity of management, timber harvest scheduling, and environmental protection.

Private owners. Nearly three-fourths (360 million acres or 140 million hectares) of all commercial forest land in the United States is privately owned by farmers, forest investment groups, other types of nonindustrial owners, or industrial firms engaged in the business of growing and harvesting timber for conversion to wood products. The objectives and practices of private owners are extremely diverse. A small, nonindustrial owner may practice extensive forestry, harvest timber infrequently, emphasize recreational activity, and treat forest assets as an immediate source of cash. A large public corporation has a financial responsibility to maximize the value of the forest and any associated facilities. Monetary objectives are therefore important in making forest management decisions: Timber is commonly harvested annually on a sustained yield basis; harvested sites are planted with genetically improved trees; competition may be removed; and growing stands may be fertilized and thinned periodically. Recreational activity such as hunting and fishing may be permitted and generate significant income. Unpriced resources may be of secondary importance to the owner.

Many states and local governments have enacted laws that regulate the practice of private forestry. Historically, the primary objectives were to assure long-term timber supply, protect streams, control soil erosion, minimize wildfire, and provide for compatible land uses. Since the 1980s, wildlife habitat and biological diversity have become important regulatory objectives in some states. Therefore, management planning for a specific property requires a detailed review of the owner's objectives, resources, and any legal constraints regarding land uses or choice of management practices in the local area.

Public agencies. One-fourth (about 120 million acres or 49 million hectares) of the commercial forest in the United States is administered by federal and state agencies. The National Forest System, managed by the U.S. Forest Service, is particularly important in the western United States, where it includes nearly one-half of the commercial forest. Congress requires that federal lands be managed for multiple uses under long-term sustained yield principles. Each national forest is required by federal law to develop a long-term land management plan, involve the public in evaluating alternatives, project future practices and outputs, and identify methods to mitigate adverse environmental impacts. These plans provide for timber harvesting, wilderness management, watershed protection, wildlife habitat, and other services in combinations that vary from forest to forest. The recovery plans of many endangered species occur primarily on public forest lands.

Land-use plans. Planning is a process for making decisions about objectives, allocating land among alternative uses, selecting management practices, determining timber harvest levels, and providing for environmental protection. Computer-based mathematical models can be used to project and evaluate the potential outcomes of alternative programs. However, the decision-makers must exercise considerable professional judgment because of uncertainty, necessary assumptions, and because resource values, for example those associated with wildlife and scenery, are not readily quantified.

A forest land-use plan is the outcome of a planning process. The complexity of the process and the plan itself will depend upon the owner's situation. Farmers and other small, private forest owners may retain professional foresters, such as extension foresters, company foresters, or private consultants, to develop plan. They may also use intuitive, informal methods. Few develop formal written plans. In contrast, federal and state agencies are required by law to develop written land management plans. Many large industrial firms also develop formal plans. Written plans vary in degree of detail and resource use issues. Generally, the ideal plan includes two parts: a map that illustrates the location of land uses, access system, and site development; and a written narrative that establishes objectives and constraints, describes and evaluates alternatives, estimates the expected flow of services including timber harvests, describes and schedules the management activities needed to implement the plan, compares costs and benefits, and discusses environmental considerations. *See* LAND-USE PLANNING.

Resource assessment. Planning and project implementation requires information about the physical, vegetative, and developmental characteristics of forest resources within the management unit. Aerial photography, satellite imagery, and statistically designed ground surveys are commonly used to obtain the necessary information. To facilitate development of management prescriptions, estimation of timber growth and yield, and environmental analyses, the unit should be divided into relatively homogeneous land classes based on such attributes as soil type, vegetation cover type, erosion hazard, slope, and age- and size-class of forest stands. Maps showing the geographical distribution of each class, existing road access, any legally restricted land uses, and boundaries of the planning unit are essential. Computer-operated software for preparing such maps is widely used. *See* GEOGRAPHIC INFORMATION SYSTEMS; REMOTE SENSING.

The resource assessment should normally include estimates of timber volume classified according to species, age or size class, quality, location, and other attributes that affect value in the local market or have relevance to decision-making. Typically, statistically designed sampling procedures using plots or strips are employed by specially trained technicians to estimate volume. Periodic surveys at 5-, 10-, or 15-year intervals provide information for monitoring changes in inventory, timber growth, regeneration following harvest, and effects of logging on environmental conditions. *See* FOREST MEASUREMENT.

Information about nontimber resources such as wildlife, streams and lakes, fisheries, and historical and archeological sites may also be required, depending upon the owner's objectives and local forest practice regulations. The identification of species classified legally as threatened or endangered is critically important to the manager, as their presence may affect land use and choice of management prescriptions. Assessment methods normally utilize both ground and aerial surveys, and professional specialists employed by the owner or by outside consultants.

Forest structure, growth, and yield. For timber management purposes, stands of trees may be classified as either even aged or uneven aged. The classification determines the method for regulating growth and yield.

In even-aged stands, all trees were established at about the same time, whether by natural (for example, germination of seeds following destruction of the previous stand by fire) or artificial means (planting of trees or seeding following a commercial timber harvest). As the stand ages, competition may eventually produce a mix of tree sizes. An even-aged stand has a life cycle with a definite beginning and ending. The period of time between these points is called the rotation. Stands which are being managed for biomass production (for example, wood to be used for the production of paper pulp), may be planted at a high density, receive no thinning, and be harvested at a point near to the stand's maximum sustainable yield. Intensively managed hybrid poplar (*Populas*) or willow (*Salix*) stands may have rotations as short as 7 years while low productivity Jack Pine (*Pinus banksiana*) stands may use a 50-year rotation. Stands which are managed for solid wood production (for example, lumber or veneer) will be maintained at a lower density in order to target growth, receive one or more thinnings, and have a longer rotation (as short as 20 years in the U.S. South to over 100 years for quality hardwoods). The sequence of practices in the second and subsequent cycles may differ from the first if conditions, including the manager's objectives, have changed.

A forest can be structured to produce a regular periodic or annual yield of timber by developing and maintaining a series of even-aged stands growing toward harvest age in a regular sequence. For example, a 10,000-acre (4050-ha) forest that is regulated under a 50-year rotation to produce an annual harvest would have 200 acres (80 ha) in each 1-year age class. Each year, the oldest age class is harvested and immediately regenerated. Theoretically, over time, total growing stock (inventory) and timber harvest remain constant. Sustainable annual harvest $H = A/t \times Y(t)$, where A is the forest area in acres, t is rotation in years, and $Y(t)$ is yield of timber per acre at age t.

Uneven-aged stands typically have a mix of trees of different ages, sizes, and species. Such stands do not normally have a specific beginning or ending because the site is continuously stocked with growing trees. Uneven-aged stands may be managed in a continuing series of partial cuttings. The period between such harvests is the cutting cycle. During each cycle, a portion of the inventory is harvested, leaving a reserve stock of trees to grow over the next cycle.

Regulation of timber harvest under an uneven-aged management system is based on the cutting cycle in a manner similar to the use of a rotation period in the even-aged system. Each acre in a 10,000-acre (4050-ha) forest managed under a 1-year cutting cycle would be partially cut each year. In practice, this may not be feasible because one-year's growth in per-acre units is usually too small for economically efficient harvesting. Cutting cycles of 10–20 years are more commonly followed. To achieve a regular annual yield, the forest area is divided by the length of the cutting cycle to determine the number of cutting units. For example, for a forest of 10,000 acres (4050 ha) managed under a 10-year cutting cycle, there would be 10 cutting units of 1000 acres (405 ha) each. Each year, one unit is selectively harvested; over the 10-year cutting cycle, each unit is harvested once. The sequence may be continued perpetually, producing a regular annual yield and maintaining reserve growing stock at a constant level.

In practice, because the demand for timber fluctuates in response to the general business cycle, under either the even- or uneven-aged management systems it may be necessary, or financially desirable, to cut more in some years and less in others. *See* FOREST TIMBER RESOURCES; REFORESTATION; SILVICULTURE.

Management scheduling. A forest typically will have complex structure. There may be a range of soil types, slopes, and aspects that differ in potential productivity. The distribution of stands and trees according to age and size classes may be irregular rather than balanced. Some stands may be understocked, while others are too dense for optimal growth. Federal or state laws may require preservation of habitat for an endangered species or stream protection.

To facilitate planning for such complex situations, optimization methods may be used to schedule management activities over time, determine the timber harvest level, allocate the land base among alternative uses, and calculate benefits and costs. A common method, linear programming, requires that the manager specify both a linear objective function to be maximized and a set of linear equations that describe management constraints. The constraints normally will include land area stratified according to vegetation-soil characteristics, beginning inventory, timber harvest-flow policy, regeneration requirements, and environmental protection requirements. Computer-based software is available to solve such problems, which typically include an automated procedure for generating the system of equations that defines the manager's objectives, alternatives, and constraints; a linear programming package; and a report writer that prepares the results in a form useful to the manager. *See* LINEAR PROGRAMMING.

Multiple use. Modern forest management usually involves the production of multiple services of value to the owner or to society. Determining the best mix of services requires technical information on trade-offs between the different outputs, costs and values, and pertinent legal constraints. Subjective evaluations may be required in the case of unpriced services such as wildlife, water, and scenery.

Forest structure has a major impact on the mix of outputs. On a forest of fixed size managed on an even-aged basis, increasing the planning rotation beyond a certain age may reduce the level of sustainable timber production but provide for a higher level of biological diversity. The populations of some wildlife species may be favored by short rotations, while others require old-growth-like conditions associated with long rotations. The technical relationships that define such trade-offs are often not available or are poorly understood. Thus, a great deal of subjective judgment by the manager is required in making multiple-use decisions.

Forest outputs or services can generally be classified as complementary or competitive. Services are complementary if an increase in the output of one is accompanied by an increase in the output of the other. They are competitive if effort to increase the output of one results in a reduction in the output of the other. Services may be complementary over one range of outputs and competitive over another.

In response to concerns about landscapes, such as maximum opening sizes, spatial diversity, and wildlife habitat, new scheduling methods are being developed and used (for example, tabu search, simulated annealing, and heuristics). Unlike linear programming, these methods do not guarantee optimal solutions. However, they do provide quality solutions that consider the complexity of modern forest management and can be implemented on the ground. *See* FOREST AND FORESTRY.

Dennis E. Teeguarden

Bibliography. M. D. Bowes and J. V. Krutilla, *Multiple-Use Management: The Economics of Public Forest Lands*, 1989; F. W. Cubbage, J. O'Laughlin, and L. S. Bullock III, *Forest Resource Policy*, 1993; L. S. Davis and K. Norman Johnson, *Forest Management*, 4th ed., 2000; W. A. Duerr, *Introduction to Forest Resource Economics*, 1993; W. A. Duerr et al., *Forest Resource Management: Decision-Making Principles and Cases*, 2d ed., 1987; W. D. Klemperer, *Forest Resource Economics and Finance*, 1996.

Forest measurement

The science and practice of measuring the volume, growth, and development of trees individually and collectively and estimating the products obtainable from them. Foresters use quantitative sciences such as mathematics and statistics for these measurements.

Regardless of the land management objectives—timber, wildlife, recreation, watershed, or a combination of these resources—the tree overstory (the forest canopy) must be quantified for informed decision making. Forest cover is an important part of wildlife habitat, and the understory component is related to the overstory characteristics. The recreation potential of wildland is a function of many variables, including the size and number of trees present. Water yields are related to the composition and density of the tree canopy. The measurement principles discussed here are applicable to all forest resource management situations that require quantitative information about the tree component of the land base.

Primary product measurement. Products cut from trees are usually measured in different units, depending on local customs and end use. The most common units are cubic volume, stacked wood measure, board-feet, and weight. Tree sections 8 ft (2.4 m) or more in length are called logs; shorter pieces are referred to as sticks or bolts.

Cubic volume. Cubic volumes of logs or bolts are computed from the length and the cross-sectional area, which may be taken at the midpoint of the log or averaged from the areas of the ends or from the weighted average of the small ends and the midsection. Although tree cross sections rarely form true circles, they are assumed to be circular for purposes of computing cross-sectional areas. Areas and volumes are computed in U.S. Customary units (square feet or cubic feet) or in metric units (square meters or cubic meters).

Stacked wood. Pulpwood and firewood are commonly sold by stacked wood measure. Stacked wood is generally measured in terms of cords. A standard cord of wood measures 4 ft × 4 ft × 8 ft (1.2 m × 1.2 m × 2.4 m); its volume is 128 ft^3 (3.6 m^3) of wood, bark, and air space. The cord content of any stack of roundwood is given by Eqs. (1).

$$\text{Cord content} = \frac{\text{width (ft)} \times \text{height (ft)} \times \text{stick length (ft)}}{128} \qquad (1a)$$

$$\text{Cord content} = \frac{\text{width (m)} \times \text{height (m)} \times \text{stick length (m)}}{3.6} \qquad (1b)$$

Cordwood customers are primarily interested in the solid-wood volume rather than the total space occupied. The solid content of a stack of wood varies with species, method of piling, diameter, length and straightness of sticks, and length of branch stubs.

Board-foot log rules. Logs that will be converted into lumber are generally scaled for their estimated board-foot content. Scaling involves determining the log diameter inside the bark at the small end and the log scaling length (in multiples of 1 or 2 ft; or 0.3 or 0.6 m). The log scaling dimensions are then used to enter a log rule, which is a table or formula showing the estimated volume, usually in board-feet, for logs of various diameters and lengths. The board-foot is equivalent to a plank 1 in. (2.5 cm) thick and 1 ft (30 cm) square; it contains 144 in.^{3} (2360 cm^3) of wood. Although the board-foot has been useful for measuring sawed lumber, it is an inconsistent unit for log scaling.

The board-foot content of sawed lumber is given by Eqs. (2).

$$\text{Bd-ft} = \frac{\text{thickness (in.)} \times \text{width (in.)} \times \text{length (ft)}}{12} \qquad (2a)$$

$$\text{Bd-ft} = \frac{\text{thickness (cm)} \times \text{width (cm)} \times \text{length (m)}}{23.6} \qquad (2b)$$

In the United States, three log rules are commonly used: the Scribner log rule, the Doyle log rule, and the International log rule. The Scribner log rule was derived by diagramming 1-in. (2.5-cm) boards drawn to scale within circles of various sizes that represent the small end of logs to be scaled. Allowance was not made for taper, and so the rule ignores all volume outside scaling cylinders projected from small ends of logs.

The Doyle log rule is based on a mathematical equation. This rule tends to underestimate small logs and overestimate large ones, but it nevertheless remains in wide use, particularly in the southern United States.

The International log rule, based on a reasonably accurate formula, includes a fixed taper allowance of $^1/_2$ in. (1.27 cm) per 4 ft (1.2 m) of log length. The International $^1/_8$-in. rule allows for saw kerf of $^1/_8$ in. (0.32 cm), whereas the International $^1/_4$-in. rule allows for saw kerf of $^1/_4$ in. (0.64 cm). The Scribner,

TABLE 1. Board-foot log rules for 16-ft (4.9-m) logs

Log diameter, in. (cm)	Scribner, bd-ft (m^3)	Doyle, bd-ft (m^3)	International $^1/_4$ in., bd-ft (m^3)
8 (20.3)	32 (0.076)	16 (0.038)	40 (0.094)
12 (30.5)	79 (0.186)	64 (0.151)	95 (0.224)
16 (40.6)	159 (0.375)	144 (0.340)	180 (0.425)
20 (50.8)	280 (0.661)	256 (0.604)	290 (0.684)
24 (61.0)	404 (0.953)	400 (0.944)	425 (1.003)
28 (71.1)	582 (1.373)	576 (1.359)	585 (1.380)
32 (81.3)	736 (1.737)	784 (1.850)	770 (1.817)
36 (91.4)	923 (2.178)	1024 (2.416)	980 (2.312)
40 (101.6)	1204 (2.841)	1296 (3.058)	1220 (2.879)

Doyle, and International $^1/_4$-in. log rules for 16-ft (4.9-m) logs are compared in **Table 1.**

As noted in Table 1, the different log rules provide varying estimates of board-foot content. Further, one would not expect any given log rule estimate to coincide exactly with the mill tally (actual lumber volume sawed). If the lumber tally is greater than the log scale, the excess is called overrun. When log scale values are larger than sawed output, the deficit is termed underrun. The overrun or underrun can be expressed by Eq. (3).

$$\text{Overrun or underrun (percent)} = \frac{\text{mill tally} - \text{log scale}}{\text{log scale}} \times 100 \qquad (3)$$

Weight scaling. Much of the pulp and paper industry buys pulp wood by weight scaling rather than linear measure. Price/weight equivalents are usually established by making thousands of paired weighings and stick scalings of loads of freshly cut wood. A weight-per-cord conversion is stated, and prices are quoted on a per-cord basis.

Sawlogs are also bought by weight, with suitable adjustments made to reflect the fact that yields of usable lumber per unit weight are less in smaller or defective logs. *See* LOGGING; PAPER; WOOD PRODUCTS.

Measuring standing trees. Standing trees are commonly measured for diameter, height, and age. The diameter and height measurements are used to estimate the volume (or weight) and value of individual trees; ages are used in assessing site quality and predicting growth.

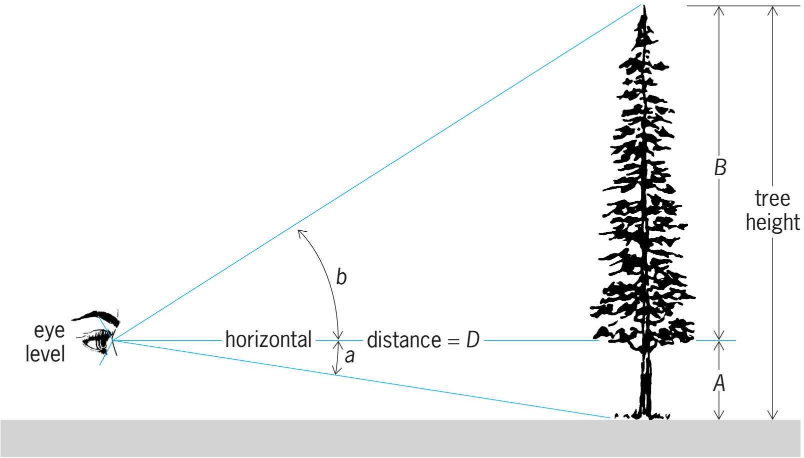


Fig. 1. Determination of tree height using common sighting instruments. *A* = tan angle *a* × *D*. *B* = tan angle *b* × *D*. Tree height = *A* + *B*. (*After T. E. Avery and H. E. Burkhart, Forest Measurement, 4th ed., McGraw-Hill, 1994*)

Diameters. The tree measurement made most frequently by foresters is diameter at breast height (dbh), defined in the United States as the stem diameter outside the bark at a point 4.5 ft (1.3 m) above the average ground line on the uphill side of the tree. The instruments used to measure tree diameters are called dendrometers. A diameter tape actually measures circumference but is graduated to provide direct readings of tree diameter. Calipers are convenient for measuring small trees. The Biltmore stick is a straight wooden slat, like a ruler, but with a specially graduated scale for direct readings of dbh.

Upper-stem diameters are occasionally obtained directly with the aid of ladders or climbing irons to reach the point of measurement, but this is time-consuming and may be hazardous. Upper-stem dendrometers have been developed to allow measurement to be made from the ground. These devices include calipers on a pole, stadia-based instruments, split-image rangefinders (such as the Barr and Stroud optical dendrometer), and the pentaprism caliper.

Heights. Height-measuring devices are called hypsometers and usually embody the basic trigonometry principle illustrated in **Fig. 1.** The observer stands at a fixed horizontal distance from the tree. Tangents of angles to the top and base of the tree are multiplied by the predetermined horizontal distance to give the height of each measured section of the tree. For accurate results, trees must not lean more than 5° from the vertical, and the fixed horizontal distance must be carefully determined. *See* SURVEYING.

For less precise work, a Merritt hypsometer may be used. This has a linear scale, often printed on one face of a Biltmore stick and based on the principle of similar triangles. It is normally used for determining merchantable log heights rather than total tree heights.

For smaller trees, direct linear measurement of tree height may be made by using height poles, which consist of joined tubing with colored stripes at regular (1- or 2-ft, or 0.3- or 0.6-m) intervals.

Foresters generally measure either total tree height or merchantable height, although in some instances both are measured. Total tree height is the linear distance from ground level to the upper tip of the tree crown. The tip of the crown is easily discerned when trees have conical shapes, but it may not be readily determined for deciduous trees with irregular

TABLE 2. Merchantable volume* of ponderosa pine trees, in cubic feet (cubic meters)

Diameter at breast height, in. (cm)		Total height above ground	
	40 ft (12 m)	60 ft (18 m)	80 ft (24 m)
6 (15.2)	3.2 (0.091)		
8 (20.3)	5.8 (0.164)	8.9 (0.252)	
10 (25.4)	9.1 (0.258)	13.8 (0.391)	
12 (30.5)	13.1 (0.371)	19.8 (0.561)	26.6 (0.753)
14 (35.6)	17.7 (0.501)	26.8 (0.759)	35.9 (1.017)
16 (40.6)	23.1 (0.654)	34.8 (0.985)	46.6 (1.320)
18 (45.7)	29.1 (0.824)	43.9 (1.243)	58.7 (1.662)
20 (50.8)	35.9 (1.017)	54.0 (1.529)	72.2 (2.044)
22 (55.9)	43.3 (1.226)	65.2 (1.846)	87.0 (2.464)
24 (61.0)		77.3 (2.189)	103.3 (2.925)
26 (66.0)		90.6 (2.565)	120.9 (3.423)
28 (71.1)		104.8 (2.968)	139.9 (3.961)
30 (76.2)		120.1 (3.401)	160.3 (4.539)

*Inside-bark volume, from a stump height of 1 ft (0.3 m) to a 4-in. (10-cm) top diameter.

or round-topped crowns. Consequently, total height is most often applied when measuring coniferous trees.

Merchantable tree height refers to the usable portion of the tree stem, and is defined for single-stemmed trees as the length from an assumed stump height to an arbitrary upper-stem diameter. If the main trunk does not extend as a single stem to the tree tip, the upper limit of merchantability is generally dictated by the branching characteristics.

Age. Many tree species found in the northern temperate zone grow in diameter by adding each year a single, distinctive layer of wood. This layer consists of a band of early or spring wood and a band of less porous, darker late or summer wood. The two bands make up an annual ring, which appears in a cross section of the stem as a series of concentric rings that can be counted to give the age of the tree.

Complete cross sections give the most reliable ring counts but are often unavailable. Ages of standing trees are therefore often determined by extracting a radial core of wood with an instrument called an increment borer. The reliability of annual ring counts depends on the species and the growth conditions. Prolonged drought or defoliation by insects followed by favorable growing conditions can create "false rings," which are often incomplete and distinguishable only in full cross sections. Heavily suppressed forest trees can also have cross sections with more than one year's growth represented by an annual ring. Such "missing rings" may be evident along only part of the tree length. *See* DENDROLOGY; TREE GROWTH; WOOD ANATOMY.

Volumes and weights. A volume table gives the average contents for standing trees of various sizes. The most commonly employed units of volume are board-feet, cubic feet, cubic meters, and cords. Volumes may be listed for some specified merchantable portion of the stem or for the total stem. In modern practice, equations are generally used to predict tree volumes rather than obtaining the values from tables. However, volume table has persisted in forestry usage as a generic term meaning tabulations or equations that show the contents of standing trees. Tree weight tables are analogous to volume tables except that weights (green or dry) of standing trees are predicted rather than volumes.

Volume tables based on diameter at breast height and tree height are referred to as standard or multiple-entry tables (**Table 2**). Local or single-entry volume or weight tables are based solely on diameter at breast height.

The most commonly used equation for deriving standard (multiple-entry) volume or weight tables is the so-called combined variable equation (4), where

$$V = b_0 + b_1 D^2 H \tag{4}$$

V is the tree volume or weight, D the diameter at breast height, and H a measure of total or merchantable tree height. The constants b_0 and b_1 are based on sample data.

Tarif tables. Among local or single-entry volume tables, tarifs are of special interest. Composed of tabulated information, they are based on the assumption of proportionality between the volume of a tree and its basal area (the cross-sectional area at the diameter at breast height). The constant of proportionality varies with the species, age, height, and other factors. Tarif numbers are averaged from height and diameter-at-breast-height data of a stand of trees and used to construct a local volume table based solely on diameter at breast height.

Biomass estimation. With increasing emphasis on complete tree utilization and use of wood as an energy source, methods have been developed for determining the weights of trees and their components (bole wood, bark, branches, foliage). The results are expressed in terms of dry weight and incorporated in biomass tables, which may include only the aboveground portion or the entire tree, including roots. Like tree volume and weight, biomass is estimated from diameter at breast height and tree height.

Timber inventory. The purpose of a timber inventory, or timber cruise, is to estimate the volume (or value) of standing trees in a given area. This is usually a sampling process. Other things being equal, the intensity of sampling tends to increase as the size of the tract decreases and as the value of the timber increases. Foresters commonly conduct timber inventories for land acquisition, timber sales, management planning, and special assessments, such as evaluating regeneration and insect or disease outbreaks.

Sampling methods vary. They may, for example, use continuous strips of uniform width spaced uniformly through the forest (**Fig. 2**). Another technique is to use a set of fixed-area plots located throughout the area of interest. The sample plots can be located by any random or systematic sampling design. However, the most common procedure is to space the plots at intervals on lines separated by a fixed distance (**Fig. 3**). This is termed a line-plot cruise.

Forest monitoring. In addition to inventories aimed at determining current conditions, land managers need trend data. Monitoring consists of collecting information over time, generally on a sample basis by measuring change in key indicator variables, in order to determine the effects of management treatments in the long term. These data, along with research results, can be used to modify management on a continuing basis to ensure that objectives are being met. The sampling design for monitoring generally involves repeated measurements on the same sample plots or individuals.

Geographic information systems. Forest inventory information is commonly stored, updated, and retrieved through geographic information systems (GISs). A GIS is a computerized database for storing, manipulating, and displaying map (spatial) data and tabular (attribute) information. In a GIS, forest inventory information can be stored in a computer and directly linked to associated forest maps, making it easier and faster to analyze and graphically display the results of forest inventories.

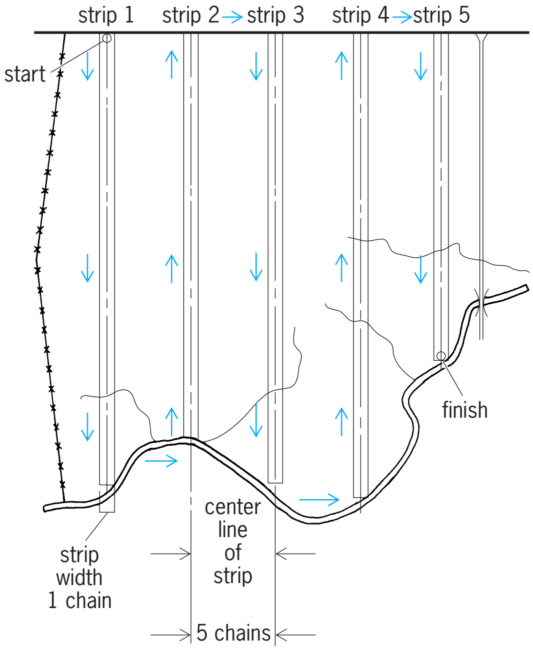


Fig. 2. Plan for a nominal 20% systematic strip inventory. Sample strips 1 chain (66 ft or 20 m) wide are spaced 5 chains (330 ft or 100 m) apart. (*After H. E. Burkhart et al., Timber inventory*, *in K. F. Wenger, ed., Forestry Handbook, 2d ed., John Wiley, 1984*)

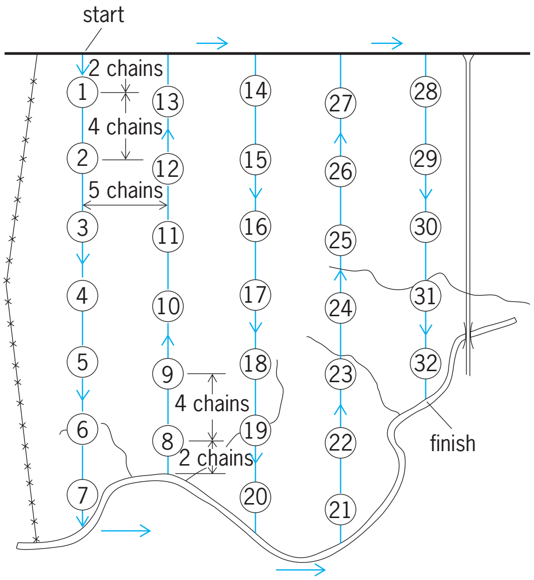


Fig. 3. Plan for nominal 10% systematic plot inventory. Sample plots are circular, $^1/_5$ acre (0.08 ha) in size, spaced at 4-chain (264-ft or 80-m) intervals on lines 5 chains (330 ft or 100 m) apart. (*After H. E. Burkhart et al., Timber inventory*, *in K. F. Wenger, ed., Forestry Handbook, 2d ed., John Wiley, 1984*)

GISs can make forest inventory information more powerful by allowing forest resource managers to integrate it with other data commonly needed to make management decisions. For instance, managers may need to know the location of roads, streams, or sensitive soils when developing management plans. Combining forest inventory data with other land resource information allows managers to make more informed decisions.

In a GIS, the basic types of data are separated into individual maps called themes (**Fig. 4**). The locations of each feature in each theme are stored in the computer database, and the characteristics associated with that feature are linked to it. For example, the locations of stand boundaries are stored in the stands theme of the GIS, along with the inventory results (for example, number of trees, volume) for each stand. Similarly, the locations and attributes (such as the flow rate, pH) of streams, soil types with their characteristics (such as, soil series, drainage class, texture, erodability class), and other relevant resources are entered into the GIS. Each of these themes can be examined individually or in combination with others.

Site quality. Foresters estimate site quality to assess present and future forest productivity and to provide a frame of reference for land management. Many parameters that affect productivity are difficult or impossible to measure directly, and as a consequence site quality is determined indirectly.

TABLE 3. Yield table for thinned stands of yellow-poplar, in total cubic-foot volume per acre (m^3/ha)*

Age, years	Site index,† ft (m)	Basal area 50 ft^2/acre (11.5 m^2/ha)	Basal area 90 ft^2/acre (20.7 m^2/ha)	Basal area 130 ft^2/acre (29.8 m^2/ha
20	90 (27.4)	1032 (72.2)		
	110 (33.5)	1266 (88.6)		
	130 (39.6)	1459 (102.1)	2582 (180.7)	
40	90 (27.4)	1789 (125.2)	3166 (221.5)	
	110 (33.5)	2195 (153.6)	3884 (271.8)	5552 (388.5)
	130 (39.6)	2529 (177.0)	4475 (313.1)	6395 (447.5)
60	90 (27.4)	2149 (150.4)	3804 (266.2)	5436 (380.4)
	110 (33.5)	2637 (184.5)	4666 (326.5)	6669 (466.6)
	130 (39.6)	3037 (212.5)	5375 (376.1)	7682 (537.5)
80	90 (27.4)	2356 (164.9)	4169 (291.7)	5958 (416.9)
	110 (33.5)	2890 (202.2)	5114 (357.8)	7309 (511.4)
	130 (39.6)	3329 (232.9)	5891 (412.2)	8420 (589.2)

*Yields are for wood and bark of the total stem for all trees with a diameter at breast height of 4.5 in. (11.4 cm) and larger.
† Base age 50 years.

Most commonly, site quality is evaluated from tree height in relation to age. Theoretically, height growth is sensitive to differences in site quality, little affected by varying stand-density levels, relatively stable under varying thinning intensities, and strongly correlated with volume. Thus height has been found to be a practical, consistent, and useful indication of site quality. Site index is defined as the average height of the dominant trees (those in the upper portion of the main canopy) at an arbitrary index age (typically 25, 50, 100 years). As generally applied, site index is estimated by determining the average total height and age of dominant and codominant trees in well-stocked, even-aged stands. When these two variables (height and age) have been ascertained for a given species, they are used as coordinates for interpolating site index (height at index age) from a specially prepared set of curves, or for substitution into a site index equation.

Stand density and growth. Quantitative measures of stand density are used when deriving silvicultural prescriptions and predicting growth and yield. The two most commonly used measures of stand density are tree basal area per unit area and number of trees per unit area. Stand basal area is the cross-sectional area at diameter at breast height of all stems, or some specified portion of the stand, expressed on a per-acre (or per-hectare) basis. Similarly, trees per acre may be determined for all stems or for some specified portion of the stand. *See* SILVICULTURE.

Although growth rates vary widely among different tree species, the general pattern of growth is remarkably consistent. When expressed as a cumulative function, all measures of growth follow a characteristic sigmoid curve like that shown for height in **Fig. 5**.

Forest management decisions require information about both current and future resource conditions. Inventories taken at one instant in time provide data on current volumes and related forest characteristics. However, forests are dynamic biological systems that are continuously changing, and information on change is also needed. Stand dynamics (that is, growth, mortality, and reproduction and related changes in the stand) are predicted with growth and

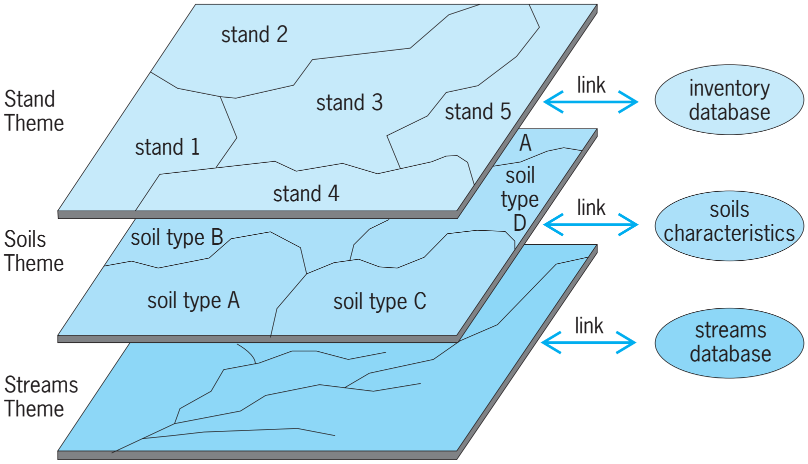


Fig. 4. Data from diverse sources can be stored, manipulated, and displayed by use of geographic information systems. (*After T. E. Avery and H. E. Burkhart, Forest Measurements, 4th ed. McGraw-Hill, 1994*)

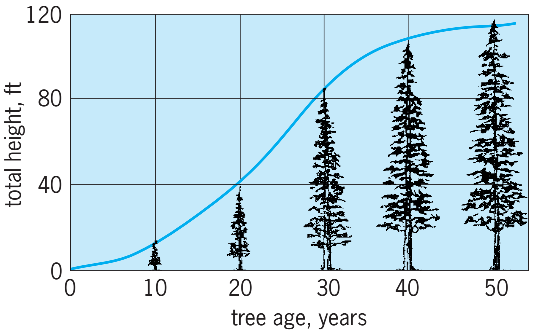


Fig. 5. The characteristic sigmoid growth curve of trees is illustrated by the cumulative height-growth pattern of some coniferous species (1 ft = 0.3 m). (*After T. E. Avery and H. E. Burkhart, Forest Measurements, 4th ed., McGraw-Hill, 1994*)

yield models. These models are quantitative representations of complex biological processes.

Growth is the increase (increment) over a given period of time. Yield is the total amount available for harvest at a given time, that is, the summation of the annual increments. The factors most closely related to growth and yield of forest stands of a given species composition are the point in time in stand development, the site quality, and the degree to which the site is occupied. For stands of uniform age, these factors can be expressed quantitatively by stand age, site index, and stand density, respectively. For a given site index and initial stand-density level, volume per unit area (yield) plotted over stand age results in a sigmoid curve. Yields for thinned stands of yellow-poplar are shown in **Table 3**. *See* FOREST AND FORESTRY; FOREST MANAGEMENT. Harold E. Burkhart

Bibliography. T. E. Avery and H. E. Burkhart, *Forest Measurements*, 4th ed., McGraw-Hill, New York, 1994; J. F. Bell and J. R. Dilworth, *Log Scaling and Timber Cruising*, O.S.U. Book Stores, Corvallis, Oregon, 1997; M. S. Philip, *Measuring Trees and Forests*, 2d ed., CAB International, Wallingford, UK, 1994; H. T. Schreuder, T. G. Gregoire, and G. B. Wood, *Sampling Methods for Multiresource Forest Inventory*, John Wiley, New York, 1993; K. F. Wenger (ed.), *Forestry Handbook*, 2d ed., John Wiley, New York, 1984.

Forest recreation

The use of forests for recreation is increasing. For many people, outdoor recreation involves direct contact with forests in various activities ranging from walking in the woods to wilderness backpacking. The primary suppliers of forest recreation opportunities in the United States are federal land management agencies such as the U.S. Department of Agriculture (USDA) Forest Service, the National Park Service, and the Bureau of Land Management; state recreation, park, wildlife, and forest departments; and private landowners and corporations (ski areas, industrial forests, resorts). The federal land management agencies administer over 650 million acres (265 million hectares) open for outdoor recreation, while there are more than 11 million acres (4.5 million hectares) of state lands and over 180 million acres (73 million hectares) of private lands open or leased for outdoor recreation.

About 95% of all Americans 16 years and older participated in outdoor recreation activities in 1994–1995, and participation has been increasing for some time. In 1960, for example, 13 million people 12 years and older participated in camping, and by 1995 there were 58.5 million campers. Over the same period, participation in snow skiing increased from 2.6 million to 26.2 million. Projections of participation in many activities for the next 50 years suggest continued dramatic growth in forest recreation. Outdoor recreation has become so popular in certain areas that some national parks and forest areas have weekend use within their boundaries that equals the population of small cities.

Management. Such growth in outdoor recreation participation requires fairly sophisticated management attention and techniques. In the early decades of the twentieth century, recreation management was mostly custodial management, keeping areas free from litter and pollutants and providing fire protection. As recreation use increased dramatically in the 1950s, a concept of activity-based management emerged with a focus on the numbers of users, the activities in which they participated, and the sites necessary for participation. In 1958 the U.S. Congress empaneled a commission to assess the situation and make recommendations for the future of outdoor recreation. It was called the Outdoor Recreation Resources Review Commission. The commission's work set the stage for significant investment in recreation facilities and areas, in recreation education and research, and in policy development focused on outdoor recreation at national and state levels. *See* FOREST MANAGEMENT.

As outdoor recreation use of forests grew and as research began to focus on behavior of participants, the activity-based management concept evolved into a more sophisticated concept of experience-based management that focused on the achievement of satisfying experiences from recreation engagement. Opportunities for experiences became the objective of management, and sites and areas are now managed to provide opportunities for a variety of experiences. Three important concepts that have come from experience-based management are a definition of recreation that is behaviorally based, the Recreation Opportunity Spectrum (ROS) approach to recreation classification and management, and the Limits of Acceptable Change (LAC) planning system. Experience-based management stems from the concept that recreation as a human experience is an intrinsically rewarding voluntary engagement during nonobligated time. The key change that this concept brought to recreation management was that attention to what recreationists experience is critical to their satisfaction. It suggests that management of places (sites and areas) and situations (numbers and behaviors of other recreationists, level of facility development) influences the experiences realized by participants.

In its most popular form, the Recreation Opportunity Spectrum visualizes the array of recreation opportunities possible, ranging from modern-urban to those described as primitive and unconfined. Most forest recreation occurs in the part of the spectrum that includes categories such as roaded-modified, roaded-natural, semiprimitive motorized and nonmotorized, and primitive, thus characterizing recreation that takes place in forest areas where environments have been modified, to recreation in wilderness and other remote and relatively unmodified areas. Using specific criteria to define different types of outdoor recreation, the Recreation Opportunity Spectrum allows managers to characterize the kind of recreation to be offered and to guide management decisions. It also is used to assess the impacts of recreation and proposed recreation on other uses such as timber management.

The Limits of Acceptable Change has replaced an earlier notion of recreation-carrying capacity that dealt with how much use is too much, and has focused attention on acceptable conditions of both the forest and the social situation. It incorporates the Recreation Opportunity Spectrum and explicitly makes monitoring and evaluation of use and impacts a part of planning and management. While the Limits of Acceptable Change has been mostly applied to wilderness recreation, it and its various agency-specific versions are applicable to virtually all forest recreation situations. The Limits of Acceptable Change is important in integrating forest recreation issues and concerns into multiple-use land management.

Quite recently, management has evolved to focus on the potential benefits that recreationists, society, and communities might realize from recreation opportunities and participation. This has become known as benefits-based management, and expected benefits are used to guide design of recreation sites and their access systems, to develop recreation information programs, and to integrate recreation with other forest uses. Benefits-based management is an outgrowth of experience-based management that extends the consideration of impacts on recreationists (the kinds of experiences they expect and realize) to impacts on communities and society as well.

Effects on ecosystems. The benefits-based management paradigm and the Recreation Opportunity Spectrum framework for recreation planning and management allow consideration of the relationship of forest recreation to ecosystem management. Ecosystem management deals with large-scale, collaborative and integrated management, and the spatial features of the Recreation Opportunity Spectrum ideally fit such a landscape view of management that cuts across ownership boundaries and vegetative and other characteristics of landscapes. It permits the assessment of the impact of recreation on the environment and on its other potential uses, as well as permitting assessment of the impacts of these other uses on potential recreation opportunities. Through its evaluative factors such as type of access, degree of management intrusion, and remoteness from sights and sounds of human activity, current recreation opportunities can be characterized, and the management actions necessary for moving toward desired recreation opportunities can be determined. Likewise, the impacts of changes in these factors due to natural or human-caused actions can be assessed for their effect on recreation opportunities.

Recreation is not neutral in its impact on ecosystems. The educational, physical, and mental health effects on individuals engaged in recreation on forest lands can have positive effects on sustainability of environmental and natural resources. Likewise, recreation can be detrimental. Wildlife can be adversely affected, as in some instances of harassment by off-road vehicle users, and of traffic on recreational road. Campsites and campgrounds can be expanded, resulting in the loss of vegetation and soil permeability. Trees and other vegetation can be chopped down for firewood. Trails and roads can be impacted by use such that they fail and lead to problems of soil erosion and stream and lake pollution. Campers can introduce soap and other pollutants into lakes and streams. The sheer volume of recreational use means that people cover much of the landscape with recreational activity, and without careful and considered use, significant impacts to the environment can occur. A major challenge for forestland managers is to help people achieve their recreational goals, but in ways that minimize negative impacts on ecosystems. Many activities, from educational programs to the design of facilities, can be effective in this work. The benefits-based management paradigm helps focus on the benefits that are to be realized and the appropriate supportive management techniques.

Future. New equipment, new activities, and changes in available time will impact outdoor recreation use in the twenty-first century. Growth is expected in participation in most activities, but for a few such as hunting, participation is expected to decrease. High-adventure activities and extreme sports are expected to increase dramatically. The use of outfitters and others to package recreation opportunities for those with limited time for planning also is likely to increase. Finally, forest recreation managers will develop ways to more fully integrate recreation with other land management activities to meet the needs of recreationists. *See* FOREST AND FORESTRY; FOREST ECOSYSTEM.

Perry J. Brown

Bibliography. H. K. Cordell, *Outdoor Recreation in American Life: A National Assessment of Demand and Supply Trends*, Sagamore Publishing, Champaign, IL, 1999; R. E. Manning, *Studies in Outdoor Recreation: Search and Research for Satisfaction*, 2d ed., Oregon State University Press, Corvallis, 1999; G. W. Sharpe et al., *Introduction to Forest and Renewable Resources*, 6th ed., McGraw-Hill, New York, 1995.

Forest soil

Weathering of underlying rock, climatic factors (such as rain,wind, and temperature fluctuations), living organisms, topography, and disturbances (such as glacial action, flooding, erosion, or agriculture) interact over time to produce soil. Soil helps sustain the productivity and diversity of terrestrial life by serving as the source of nutrients and water for plants, by physically anchoring plants, and by serving as the habitat for soil organisms. Far more numbers and species of organisms live belowground than aboveground. Because soil usually develops quite slowly, soil conservation is important for maintaining the productivity of the land. Soil also acts as a water purifier; organic substances and most pollutants are arrested in upper soil horizons, purifying water before it enters streams or aquifers. Soil is the repository of life; as plants and animals die, they become part of the soil. Soil profiles reveal much about the soil history, including plant and animal remains, disturbances such as flooding, volcanic ash deposits, windthrow events, rotting wood, charcoal from fires, and past human activities.

Soil types. Variations in the bedrock and other types of parent material, climate, types of plants, topography, and disturbance history result in a wide range of soil conditions. The primary minerals that initially constitute bedrock are created deep below the surface under high pressure and temperature. As these rocks are exposed to the atmosphere through geological processes, they become thermodynamically less stable. Physical and chemical breakdown (weathering) creates products, mostly organic matter and secondary minerals, common as soil constituents. Plants can speed weathering and soil development by producing acids and organic chelates (organic compounds that speed weathering by reacting with end products to form organomineral complexes) that degrade primary minerals. With abundant rainfall and anion carriers such as bicarbonate and nitrate ions, some products of weathering move in solution through soils to stream water.

Forest soils differ from agricultural soils mainly by history and types of disturbance and topography and climate. Plowing, which homogenizes the upper soil layers, is feasible only on nonmountainous terrain. Areas too dry to support forests are sometimes farmed by irrigating soils. Soil disturbances not found in agriculture but common in forests include windthrow, landslides, and log-yarding practices. These disturbances vary much more in time and space than do agricultural disturbances, making forest soils much more heterogeneous. Erosion can be extreme in steep terrain, especially after fire removes plant cover and heavy rain follows. In general, though, disturbances of forest soil are far less frequent than of agricultural soil, allowing distinctive horizons to develop. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Four common forest-soil horizonation patterns are spodosols (podzols or mor soils), mollisols (brown or mull soils), inceptisols, histisols (organic soils), and oxisols (weathered tropical soils), although many others are recognized by various classification systems. Spodosols are found mostly at mid to high latitudes under cool, moist climates, most often with conifers. They are characterized usually by clearly visible horizons: a thick organic (O) layer of partly to highly decomposed plant litter, a highly leached, light-colored (E) layer at the top of the mineral soil overlying several (such as Bh or Bs) layers of accumulated products of weathering, and overlying unaltered parent material (C). Acids move downward, weathering the upper E horizon and depositing the products of weathering in the B horizons. Slightly warmer conditions; deciduous plants and grasses; frequent mixing of the organic and upper mineral soil layer by earthworms, insects, and other soil animals; and other disturbances can create mollisols. Inceptisols are soils with less organic matter accumulation than in podzols or mollisols, because of climate, parent material, and mixing. Lack of mixing and very wet conditions can lead to histosols, where plants root only in deep organic horizons. Oxisols are very old soils with little erosion, found on flat terrain in the tropics, where weathering proceeds so far that even silicates are lost, leaving mainly iron and aluminum oxides. These soils can become quite infertile, although some tropical trees can root to distances of more than 30 m (100 ft) to less weathered layers.

Cycling of carbon, nutrients, and water. Carbon, nutrient elements, and water move and are chemically changed in highly interrelated and cyclical patterns. These cycles can be rapid or slow, and over small to large spatial scales. As forest plants, microbes, and animals grow, they alter the soil. In the carbon cycle, autotrophs (green plants), through photosynthesis, use solar energy to capture carbon dioxide (CO_2) from the atmosphere and convert it to sugars and carbohydrates (especially cellulose and lignin). Soil heterotrophs (fungi, bacteria, and soil fauna) live by consuming live roots or decomposing dead litter. As organic matter is decomposed, it passes through many forms, usually ending as carbon dioxide released back to the atmosphere, and some quite stable carbon compounds, referred to as humus. As organic matter in mineral soil increases, generally so does plant growth. Organic matter increases water infiltration, water-holding capacity, and available nutrients in exchange sites and reduces erosion. Thus, carbon cycling influences how fast rain can infiltrate into soil and how long it can be held there before being adsorbed by plants and then transpired back to the atmosphere.

Most nutrient elements, except nitrogen and sometimes sulfur, originate as part of primary minerals in silicate rocks. Elements are released from primary minerals through physical and chemical weathering. Physical weathering results, for example, from glaciers that grind rocks beneath them, and freeze-thaw cycles that shatter rocks. Chemical weathering is in part driven by carbon dioxide, released by respiring roots and heterotrophs in soil, and dissolves in water to form carbonic acid. Certain plants and microbes greatly speed weathering by exuding organic acids and chelates. Some of the nutrient elements released by weathering are held on exchange sites where they are available for plant uptake. Elements incorporated into plant biomass are then cycled with carbon, although many plants extract nutrients out of senescing leaves before they fall to the ground. Elements such as potassium, calcium, and magnesium tend to be exported easily with anions in solution; others such as iron and aluminum tend to be less mobile and accumulate over time. Eventually, most elements are dissolved or carried as sediment into rivers and oceans; they are deposited on the ocean bottom and eventually reformed into primary minerals.

Nitrogen differs from other nutrients by being rare in rocks and common as a gas (molecular N_2) in the atmosphere. Only certain bacteria can form a special enzyme (nitrogenase) which, with substantial consumption of carbohydrates, breaks apart N_2 to form amino acids and proteins. In nature, free-living N_2-fixing microbes and a few plants that can harbor N_2-fixing bacteria in root nodules (called symbiotic N_2-fixing plants, such as the legumes and alders) play very important roles in controlling the long-term productivity of forests limited by nitrogen supply. Because nitrate (NO_3^-) is an easily soluble anion, it moves freely in soil and can be lost by leaching when plant roots are not active. Bacteria that convert ammonium (NH_4^+) to NO_3^- [known as nitrifiers] and bacteria that convert NO_3^- back to dinitrogen (N_2) [known as denitrifiers] are important in nitrogen cycling as well.

Shrubs and minor species can have especially important effects on carbon and nutrient cycling. Shrubs, for example, tend not to be able to accumulate carbon in woody tissues as well as trees do. Therefore, more carbon fixed by photosynthesis is returned as litter. Deciduous trees also tend to return more litter to the soil than do evergreen trees. By doing so, shrubs and hardwoods speed many nutrient cycles and may help build soil organic matter. Different species result in different amounts and types of soil organic matter. Conifers tend to develop thick organic horizons and are more often associated with podzol-type soils; hardwoods are more often found on mull soils with thin litter layers. Plants that establish after soil is disturbed are often called pioneering or early-successional plants. By fixing N_2 and increasing the rate of weathering, they tend to alter soils more quickly than do late-successional plants. Other plants have adapted to wet soils and can tolerate and grow in low-O_2 soils. A few can even oxygenate their root zones.

Rates of soil formation. Soil forming on hard bedrock such as granite takes thousands of years. Lichens and bacteria slowly etch the rock surface, creating a weathering rind. Specially adapted higher plants become established in cracks, and eventually a soil is created. Soils can develop on newly exposed sand dunes, glacial soils, mudflows, and landslides remarkably fast, forming distinctive horizons in less than 100–200 years. Soils can also develop on some sedimentary bedrock quickly in the humid tropics. Disturbances such as windthrow and landslides and deep-rooting plants can speed weathering and create deeper soils over time. Erosion actually increases plant productivity in older soils by removing highly weathered layers, giving plants access to unweathered minerals and the nutrients they contain.

Soil-plant interface. Many of the carbon and nutrient cycling reactions are at the soil-root interface. Plants or microbes, to maintain ionic balance, must release a proton or hydroxyl ion into solution to be able to take up a similarly charged nutrient ion. Some plants and microbes release organic acids that act to speed weathering. Plants and microbes can also release enzymes, such as phosphatase, that can react with soil organic matter, releasing phosphorus into soil solution. Higher plants differ from lower plants and free-living bacteria and fungi by possessing the ability to extend roots deep into the soil, creating a larger surface area for nutrient and water uptake. A special symbiosis between most plants and some fungi called mycorrhizae allows plants to access an even larger surface area from which they draw nutrients and water. One form, ectomycorrhizae, creates a fungal sheath around small roots and has mycelia that extend inward between cortical root cells (where carbon substrates are provided by the plant) and outward into the soil where mycelia absorb nutrients and water. Mycelia and their hyphae are much finer than roots and can grow into 5-micrometer crevices, expanding the effective volume of soil explored by plant roots. Plants also exude carbohydrates into the soil and alter the microbial composition and activities in soil adjacent to roots, in a zone called the rhizosphere. *See* MYCORRHIZAE; ROOT (BOTANY).

Effects of forest disturbances. Disturbance of forest stands affects soils to different extents. Some stand-replacing disturbances where new soil is deposited, such as volcanoes, glaciers, flooding, and massive landslides, can create new soil surfaces. These soil surfaces may be poorly productive at first because of little organic matter and available nutrients, but can quickly become productive and even exceed productivity of older, undisturbed soils. Fires that lead to deep erosion of upper soil layers, exposing poorly weatherable bedrock such as granite, may take thousands of years to develop a new soil. Landslides exposing more weatherable sedimentary bedrock in the wet tropics can become soil in less than a hundred years. Soil mixing by windthrow, landslide, flooding, or dust deposit is likely to increase forest productivity by preventing nutrient immobilization in wet organic horizons.

Harvesting trees affects soils by reducing evapotranspiration and water flow through soil, and decreasing nutrient uptake. Nutrients, especially NO_3^-, can be quickly lost to stream water when plant roots are not active. Harvesting of stems and other tree parts also removes nutrients from the site. Short crop rotations and whole-tree harvesting techniques are likely to lead to nutrient exports exceeding the capacity of subsequent trees to obtain nutrients, and fertilization or crop rotation with N_2-fixing plants may be required to maintain growth rates. Hot crown and slash fires can lead to massive volatilization of nitrogen and other nutrients, especially if followed by heavy rains. Generally, light underburning of the litter layer removes far fewer nutrients and may even increase the availability of nutrients tied up in organic horizons. Heavy harvesting and yarding equipment may also cause compaction and erosion, which can reduce tree growth. This concern has led to elaborate cable-yarding systems to avoid ground-disturbing equipment, especially on steep slopes. Quick reestablishment of pioneer species as well as crop trees is important to reestablish biotic regulation of carbon, nutrient, and water cycles. *See* ACID RAIN; FOREST AND FORESTRY; FOREST ECOSYSTEM; SOIL; SOIL ECOLOGY. Bernard T. Bormann

Forest timber resources

The timber produced by temperate, boreal, and tropical forests around the world. Timber is included in a wide variety of industrial and other commodities that are traded in national and international markets.

United States resources. Of the total land area of the United States, forests cover one-third, that is, about 737 million acres (295 million hectares). Approximately two-thirds of the total forest area in the United States is classified as timberland, that is, forestland capable of producing at least 20 ft^3 (0.6 m^3) of industrial wood per acre per year, and not reserved for uses incompatible with timber production. The remaining forest area meets the growth criteria for timberland, but has been set aside for parks and wilderness areas or is incapable of producing a sustained crop of industrial wood.

Most nontimberland is in public ownership, including 36 million productive forest acres (15 million hectares) legally withdrawn for such uses as national parks, state parks, and national forest wilderness areas. Of the remaining nontimberland, about 107 million acres (43 million hectares) are in Alaska (part of the Pacific Coast region), and much of this acreage is boreal forest as opposed to the temperate forest types prevalent in most of the United States.

Between 1952 and 1995, the area of timberland had a net decrease of 4%. Much of the decline was the result of withdrawals of public timberland for wilderness or for other land uses that do not permit timber harvest. About half of the decline was offset by conversions of abandoned cropland to softwood timber types, especially in the southern United States. From 1985 to 1995, the timberland area remained fairly stable. There was, however, some reclassification from timberland to other forest areas as a result of reevaluation of site productivity, and some locally significant loss of timberland to urbanization. In the southern United States, the loss of timberland to agricultural uses was stemmed in part by reductions in agricultural production.

Forest types. Most timberland in the United States is in the temperate zone. The broad distinction made among temperate tree species is softwoods (the conifers, or cone-bearing trees, such as pine, fir, and spruce) versus hardwoods (the broad-leafed trees, such as maple, birch, oak, hickory, and aspen). The five largest timber types are oak-hickory, southern pine, maple-birch-beech, Douglas-fir, and oak-pine.

In the east, hardwood types comprise nearly two-thirds of the forest area. Important eastern timber types are the oak-hickory, maple-birch-beech (found throughout the New England, Middle Atlantic, and Great Lakes regions), the oak-gum-cypress (primarily in the Mississippi Delta and other southern river bottoms), and the aspen-birch of the Great Lake states (relatively short-lived species that follow logging and fires). These forests supply much of the quality hardwoods in the United States.

Softwood types dominate the western forests. Douglas-fir and true firs and spruce constitute nearly half of the western commercial timberland. The western softwood types are a major source of lumber and plywood in the United States. Nearly all the commercial forest area of coastal Alaska is of the hemlock–Sitka spruce type. Hardwoods (mostly in Washington and Oregon) occupy only a small percentage of the commercial forest area of the western United States. *See* PLYWOOD.

Growth. Growth on United States timberlands increased 55% between 1952 and 1995, being evenly divided between softwoods and hardwoods. The fastest growth rates were in the South for both hardwoods and softwoods. In 1992, growth exceeded removals in all areas, except for softwoods in the South, where harvests increased after the late 1980s to offset reduced harvests in the Pacific Coast states because of the strategies adopted to protect the habitat of the northern spotted owl.

Timber inventory. The growing-stock inventory contained a vast amount of sound wood at the beginning of 1992. Timber inventories grew at 0.6% per year between 1955 and 1995; most of this growth was in the hardwood component.

Forestland management in the United States has produced positive results in the sense of sustaining forest resources. However, there are concerns about biodiversity, lack of certain seral stages, and loss of what is publicly perceived as nonrenewable forest resources.

The average diameter of harvested timber is a key indicator of timber utilization. For example, between 1976 and 1992 the average diameter of harvested timber fell by 35% for the Pacific Coast states, 44% for the Rockies, and 18% for the South. In the northern United States, the average diameter increased over the same period by 5% for both softwoods and hardwoods.

Ownership. Nearly three-quarters of timberland is privately owned. The remainder is held by various public owners, and comprises national forests and state, county, and municipal holdings. Timberland under Native American sovereignty accounts for about 1% and is included with other private timberlands. Private owners can be distinguished as industrial, farm, or miscellaneous. Approximately 14% of the commercial forestland is in the hands of industry, the remainder being divided between miscellaneous private owners and farmers.

Some of the most productive forestlands of the United States are in private industrial holdings. These lands are generally managed to produce high volumes of wood fiber for both the solid-wood-products and pulp and paper industries. Many timberland forest industries are concentrated in the southern states.

The miscellaneous category, which holds 34% of the privately owned forestland, consists of a great variety of individuals, groups, and non-forest-products companies. These owners range from those who are interested in producing timber to those whose objectives are nontimber. Most farm forests are managed for timber production. As such, the condition of farm-owned woodlands has been a source of disappointment, discussion, and considerable action on the part of public conservation agencies.

A diverse array of public agencies hold large forest areas, the most important being the national forests. These comprise 89 million acres (36 million hectares) of commercial forestland that is managed and administrated by the U.S. Forest Service, a bureau of the Department of Agriculture. Most public holdings are managed under multiple-use principles. There has been a dramatic shift in focus away from wood as the principal product of the national forests to sustaining ecosystems where wood is just one of the outputs.

Inventories on public timberlands have been stable since 1955—largely because withdrawals for wilderness and other uses offset inventory gains attributed to increases in growth. Private inventories, however, grew by 51% over the same period. Softwood inventories are roughly evenly divided between public and private timberland owners, but the hardwood inventory is concentrated on private timberland, especially nonindustrial ownerships.

World forest resources. Forests cover just 30% of the land area of the Earth. Included are closed and open natural forest, forest plantations, and other woodland. Shrub, scrub, and brushland cover about one-fifth of the forest area.

Half of the world's forests are in Asia, Latin America, and Africa. These areas comprise predominantly developing countries, and the forests are mostly tropical. About one-quarter of the world's forests, nearly half of the tropical forests, and more than half of the wet tropical forests (rainforest and moist lowland forest) are in Latin America. Just over half of the world's forests are in temperate zones; the boreal region in the temperate zone accounts for more than one-third of the world's forests.

Between 1980 and 1990, the worldwide area of forest declined by 4%. Deforestation across all regions occurred at rates estimated to be 0.3–0.6% per year between 1975 and 1995. An increase in the area of temperate forest over the period 1980–1990 was much more than offset by a reduction in the area of tropical forest. Lowland tropical forests account for the majority of tropical forest area and tropical deforestation; however, the rate of deforestation is most severe in the dry tropical forests, where population pressure and environmental stress are combined.

About two-thirds of the world's forests are publicly owned. Although slightly more than half of temperate forests are in public ownership, nearly all tropical forests are in public or communal ownership. The United States, Japan, and the Nordic countries have relatively higher proportions of private ownership than other developed countries. Ownership patterns affect the mix of forest outputs and the type of management practices. Although public ownership is typically thought to best assure provision of nonmarket goods and intangible forest benefits, private management is recognized as efficiently providing commodity-based forest outputs.

Plantation forests. Although plantations are estimated to be little more than 3% of world's forest area, they are an important component of industrial timber production in the temperate zone, and are a source of both industrial timber and nonindustrial products in the tropical zone.

Temperate-zone plantations are predominantly coniferous species; most are managed for industrial wood products, and are the consequence of reforestation rather than afforestation. That is, these plantations generally do not expand existing forest area. Afforestation of idle crop or grazing land (as in the United States and United Kingdom) is an exception. Native species account for the majority of the plantation area in most Northern Hemisphere temperate-zone countries. Growth rates may be somewhat higher than growth rates for natural forests as a result of the use of improved planting stock or more intensive management practices. *See* REFORESTATION.

In the temperate zone of the Southern Hemisphere, most plantations are fast-growing, exotic species, with a shorter production cycle than native species. During 1985–1995, production from exotic softwood plantations in New Zealand, Chile, and Australia increased significantly; these plantations account for a small percentage of world coniferous industrial roundwood production.

India, Indonesia, and Brazil account for the majority of plantations in the tropical zone. More than half of the plantations in the tropical zone were established for purposes other than timber production; these include plantations for erosion control and for agroforestry (that is, joint production of trees and agricultural crops). Tree plantations established for annual crop production in the tropical zone (such as rubber, coconut, and oil palm) also have contributed to the supply of wood and wood products. *See* FOREST AND FORESTRY; FOREST MANAGEMENT.

Richard W. Haynes; David J. Brooks

Forestry, urban

The planning and implementation of actions to establish, protect, restore, and maintain trees and forests in cities and smaller communities. When it began in the 1960s, the field of urban forestry focused on individual trees along streets and adjacent to homes and buildings, and on groups of trees in specific spatial areas such as parks. Since then, it has evolved toward more holistic consideration of the structure, processes, and functions of urban ecosystems at a larger scale. Scientists have come to realize that trees and forests play a critical role in maintaining healthy urban ecosystems, providing ecological services such as filtering water and air pollution, reducing stormwater flows, sequestering carbon emissions, conserving energy, and reducing soil erosion, as well as providing human health benefits, recreation, esthetics, and fish and wildlife habitat. These ecological services can be translated into economic values worth billions of dollars by comparing them to the costs of achieving the same benefits with technology and human-made infrastructure. Urban trees, forests, and related vegetation are often referred to as green infrastructure in order to compare and contrast them with human-made or hard infrastructure, such as buildings and roads. *See* ECOLOGICAL COMMUNITIES; ECOLOGICAL SUCCESSION.

Urban trees and forests also provide significant social and economic benefits. By making downtown and residential areas more livable, urban trees and forests can increase real-estate values and attract businesses; in turn, this helps to create additional employment opportunities and contributes to the overall economic viability of the community. More livable urban environments strengthen the social fabric of communities, creating more opportunity for citizen involvement in community activities, greater pride and safety in neighborhoods, and conditions that lead to greater mental and physical health within the community. Engaging community groups in urban forestry efforts can also build a stronger sense of community, empower inner-city residents to improve neighborhood conditions, and promote environmental stewardship.

Just as communities depend on healthy urban forests for ecological services, urban forests depend on socially and economically healthy communities for their care. Urban areas that lose trees and forests often go into decline, becoming impoverished environmentally, socially, and economically. Where urban areas flourish, community groups have developed a sense of civic pride and the social and economic capacity to maintain healthy trees and forests.

New technological tools. New satellite remote sensing and geographic information system tools are enabling communities to assess changes in their green infrastructure and to develop plans for restoring and maintaining urban trees and forests. City planners and policymakers use the same tools to address other urban infrastructure needs, such as transportation and housing, which allows them to integrate green infrastructure information into plans for other infrastructure development.

Recent assessments that use tree cover as an indicator of ecosystem health show that urban forests are an overlooked and underappreciated community resource. Since the early 1970s, three major metropolitan areas in the United States (Seattle, Baltimore/Washington, and Atlanta) have lost one-third or more "heavy tree cover" (50% or more). These losses, and the parallel expansion of areas with "low tree cover" (less than 20%), result in billions of dollars of lost ecological services. In the 3.9-million-acre Puget Sound area, heavy tree cover dropped by 37% between 1972 and 1996, whereas low tree cover more than doubled. When translated into economic values, this loss in tree cover and increase in impervious surface area resulted in $2.4 billion in lost stormwater retention capacity and $95 million annually in the capacity of trees to filter air pollution. Cities that recognize the economic values associated with green infrastructure can slow and reverse losses in urban trees and forests by incorporating green infrastructure into urban planning and development efforts, and by investing in plans and programs to restore and maintain urban trees and forests. Incentive-based policies may encourage private investment in urban trees and forests.

Urban and rural links. As urban forestry has expanded to include larger ecosystems, spatial boundaries between urban and rural areas have begun to blur. Their ecological, social, and economic links have come into view. Watershed connections have become a key consideration, while other social and economic links related to movements of people, businesses, goods, and services have also become clearer. Rural lands, both public and private, provide an array of ecological services to nearby cities, including drinking water, recreational opportunities, fish and wildlife habitat, and agricultural and forest products.

Sprawling urban development—the effects of people and businesses moving to areas surrounding cities—is a major policy concern with significant implications for urban forestry. Because of the expanded interpretation of urban forestry, policymakers should consider the different impacts of sprawl across the landscape. Older inner-city areas, newer communities on the urban fringe, and nearby public and private rural lands face different challenges to their green infrastructure:

1. Mature parts of cities are experiencing the movement of people and public and private financial resources to rapidly growing areas. Urban forestry efforts focus on protecting the often-limited green infrastructure and restoring degraded or abandoned (such as brownfields) sites with trees and forests. Older cities look for ways to invest in rebuilding green infrastructure to improve ecosystem functioning and social and economic well-being.

2. Newer parts of cities, often suburbs or urban fringe areas, are experiencing the greatest loss of trees and forests from sprawl because of massive public and private investment in new residential neighborhoods, public infrastructure, and private businesses. Urban forestry efforts focus on retaining critical parts of existing green infrastructure. Newer cities seek planning tools and policy mechanisms to retain significant green infrastructure in development plans and actions.

3. Sprawl has major impacts on public and private rural landscapes. Individual decisions and uncoordinated actions affect these landscapes by reducing the amount of open space, curtailing opportunities for traditional farm and forest management, and

threatening overall ecosystem health. Of particular concern is forest fragmentation, caused by the conversion of private farms and forests to other land uses and the breakup of large private ownerships into multiple smaller parcels. Policy efforts focus on retaining private lands as open space for public environmental benefits and on maintaining working lands for ecological, social, and economic benefits. As urban growth nears public forest lands (generally federal or state lands), the management of these lands often changes to meet the use preferences of urban people. National forests near major metropolitan areas have come to be called "urban national forests" and generally emphasize recreational uses and the protection of ecological services, a shift from traditional timber-oriented management.

Management. Urban forests are owned by a diverse array of public and private entities, including individual homeowners, private businesses, and federal, state, and local governments. Their management and use is closely tied to the homes, buildings, transportation systems, utilities, and other infrastructure in an urban area. A key challenge is coordinating the efforts of the large number of individuals, groups, and organizations involved in urban forestry, each with its own information, resources, and objectives. *See* LAND-USE PLANNING.

Generally, city governments guide urban forestry efforts, although their direct management responsibility focuses on city-owned lands. They receive assistance from federal and state governments and provide assistance to other landowners in the city. Because many local government agencies have jurisdiction over activities that impact the urban forest, including environmental protection, transportation, housing, and public works, there is a great need for local governments to coordinate agency planning and development activities. Most cities have some narrow policies and programs focused on planting and protecting trees, but most have not recognized the ecological, social, and economic benefits of a healthy green infrastructure through policies and programs that integrate the efforts of various local agencies.

The interest of national policymakers in urban forestry and green infrastructure has grown in recent years, as demonstrated by increasing federal support and policy initiatives. In many cities, however, financial support for urban forestry has fluctuated and often declined as local governments and policymakers have dealt with a broad array of social and economic issues. Until the general public and local policymakers understand the benefits of green infrastructure and the interdependence of healthy forests and communities, urban forestry will have difficulty building support as a policy priority of city governments.

Many professionals are involved in urban forestry efforts, including foresters, arborists, landscape architects, horticulturists, engineers, planners, wildlife biologists, and hydrologists. Planning and implementing urban forestry projects in difficult city environments requires a variety of technical and people-oriented skills. Educational programs should incorporate training in policy, social sciences, and communication, as well as the natural sciences, planning, and engineering.

The involvement of community groups and citizen volunteers is an essential element of urban forestry. Many groups plant and maintain urban trees and forests. These efforts might focus on neighborhood enhancement, stream restoration, or the transformation of brownfields into urban gardens or parks. Volunteers and community groups working in partnership with agencies and other professionals can improve urban ecosystems, community livability, and social capacity.

Outlook. Three major challenges presently face urban forestry: (1) to develop better understanding of the ecological services that urban trees and forests provide, (2) to integrate green infrastructure into city governments' planning and development efforts, and (3) to stimulate public and private investment in actions to restore and maintain urban trees and forests. To achieve these goals, the general public will need to become more aware of the interdependence between urban forests and communities, and more people will need to become involved in urban forestry activities. *See* FOREST AND FORESTRY.

Gerald J. Gray; Gary A. Moll

Bibliography. *A National Research Agenda for Urban Forestry in the 1990's*, International Society of Arboriculture, Urbana, IL, 1991; J. F. Dwyer, The human dimensions or urban forest ecosystem management, *Proceedings of the 7th National Urban Forest Conference*, pp. 44-46, American Forests, Washington, DC, 1995; G. J. Gray, Sustainable communities: Exploring uran-rural linkages, *Proceedings from the 1999 National Urban Forest Conference*, pp. 86-89, American Forests, 1999; G. Moll and S. Young, *Growing Greener Cities*, Living Planet Press, Washington, DC, 1992; *Regional Ecosystem Analysis: Chesapeake Bay Region and the Baltimore-Washington Corridor*, American Forests, 1999; C. Summers, Incorporating natural systems into urban infrastructure, *Proceedings of the 7th National Urban Forest Conference*, pp. 54-56, American Forests, Washington, DC, 1995.

Forging

The plastic deformation of metals, usually at elevated temperatures, into desired shapes by compressive forces exerted through a die. Forging processes are usually classified either by the type of equipment used or by the geometry of the end product. The simplest forging operation is upsetting, which is carried out by compressing the metal between two flat parallel platens. From this simple operation, the process can be developed into more complicated geometries with the use of dies. A number of variables are involved in forging; among major ones are properties of the workpiece and die materials, temperature,

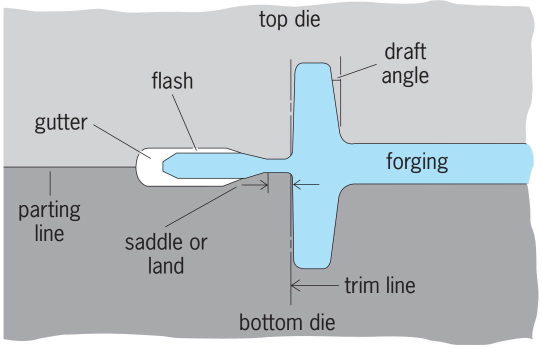


Fig. 1. Closed-die forging terminology.

friction, speed of deformation, die geometry, and dimensions of the workpiece. One basic principle in forging is the fact that the material flows in the direction of least resistance.

Friction between dies and workpiece (and, in hot forging, the more rapid cooling of the surfaces of the workpiece that are in contact with relatively cooler dies) causes barreling. Barreling can be reduced by lowering friction and also by carrying out the forging process in a more isothermal condition.

Forgeability. In practice, forgeability is related to the material's strength, ductility, and friction. Because of the great number of factors involved, no standard forgeability test has been devised. For steels, which constitute the majority of forgings, torsion tests at elevated temperatures are the most predictable; the greater the number of twists of a round rod before failure, the greater is its ability to be forged. A number of other tests, such as simple upsetting, tension, bending, and impact tests, have also been used.

In terms of factors such as ductility, strength, temperature, friction, and quality of forging, various engineering materials can be listed as follows in order of decreasing forgeability: aluminum alloys, magnesium alloys, copper alloys, carbon and low-alloy steels, stainless steels, titanium alloys, iron-base superalloys, cobalt-base superalloys, columbium alloys, tantalum alloys, molybdenum alloys, nickel-base superalloys, tungsten alloys, and beryllium. *See* METAL, MECHANICAL PROPERTIES OF.

Forging dies. Design and selection of die materials for forging involve considerable experience. Both static and impact strength are required, in addition to resistance to abrasion and to heat checking. Dies are generally heated to below 1000°F (540°C) in hot forging.

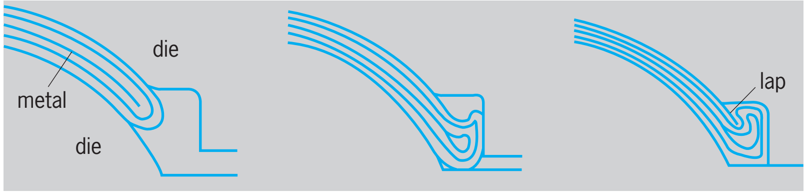


Fig. 2. Formation of a defect in closed-die forging.

Some of the terminology in forging is shown in **Fig. 1**. Draft angles facilitate the removal of the forging from the die cavity; they usually range between 3 and 10°. The purpose of the saddle or land in the flash gap is to offer resistance to the lateral flow of the material so that die filling is encouraged. Die filling increases as the ratio of land width to thickness increases up to about 5; larger ratios do not increase filling substantially and are undesirable due to increased forging loads and excessive die wear. The purpose of the gutter is to store excess metal. The flash is removed either by cold or hot trimming or by machining.

Filling of the die cavities is an important aspect and requires a careful balance of die geometry, temperature, lubrication, and speed of forging. Lubrication is important not only for reducing friction but also in its role as a parting agent, in addition to serving as a heat barrier between the die and the hotter forging. The commonest lubricant is graphite suspended in grease, oil, or water; other lubricating materials are molybdenum disulfide, talc, mica, sawdust, salt water, and glasses. Some metals, such as tungsten and molybdenum, form soft oxides upon heating, and these serve as lubricants.

Because of the many factors involved, great variations in strength and ductility have been observed in the forging of the more complex shapes. Incorrect design of dies or improper operation can result in forged products with internal defects such as laps (**Fig. 2**).

Forging design. There are in general four types of forging designs: blocker type, commercial, close tolerance, and precision. The blocker type involves the use of large radii and draft angles, smooth contours, and generous allowances. Commercial designs have more refined dimensions, such as standard draft angles of 5–7°, and smaller radii and allowances. In close-tolerance designs draft angles are in the order of 1–3°, tolerances are less than half of those for commercial designs, and there is little or no allowance for finish. Precision designs usually involve close dimensional tolerances; they require special tooling and equipment and are used mainly for aluminum and magnesium forging.

Forging equipment. A number of methods produce the necessary force and die movement for forging. Two basic categories are open-die and closed-die forging. Drop hammers supply the energy through the impact of a falling weight to which the upper die is attached. Two common types of drop hammers are the board hammer and the steam hammer. In the former the ram is attached to wooden boards which slide between two rollers; after the ram falls freely on the forging, it is raised by friction between the rotating rollers. In the steam hammer the ram is raised and lowered by a double-acting steam cylinder; the forging force is a combined effect of gravity and steam pressure.

Another type of forging equipment is the mechanical press, varieties of which are crank or eccentric, knuckle, percussion, and toggle. For large forgings the hydraulic press (with existing capacities of

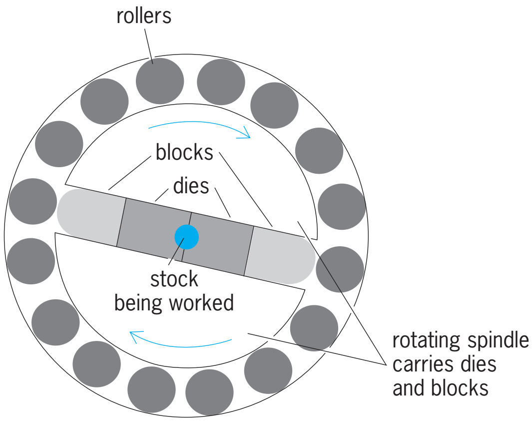


Fig. 3. Operation principle of swaging machine.

50,000 tons or 45,000 metric tons and future capacities in the order of 200,000 tons or 180,000 metric tons) is the only equipment with sufficient force. However, the speed for such presses is about one-hundredth that of hammers. Other types of forging equipment are counterblow hammers (vertical) and impacters (horizontal).

Swaging. Swaging or, as it is also called, rotary swaging consists of two or four dies which are activated radially by blocks in contact with a series of rollers (**Fig. 3**). The rotation of the block-die assembly causes the curved ends of the blocks (either circular or modified sine curve) to be in contact with the rollers; relative motion between the blocks and the roller housing then gives a reciprocating motion to the dies. The hammering action on the outer surface of the stock reduces its diameter; the stock is generally prevented from rotating.

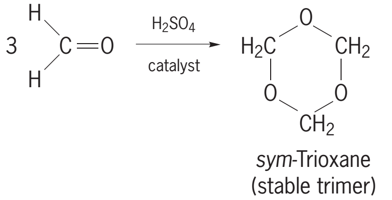
A variety of geometries is obtained in this process: (1) reduction of wall thickness of tubes; (2) sinking of tubes, whereby both the outside and inside diameters are reduced, the final wall thickness being controlled by a mandrel; (3) tapered or stepped sections; and (4) a variety of internal geometries, such as fins or hexagonal or spiral grooves, each with the use of a specially designed mandrel. Cold swaging produces excellent surface finish and dimensional accuracy, together with improved mechanical properties. *See* METAL FORMING. Serope Kalpakjian

Bibliography. B. Avitzur, *Metal Forming*, 1980; National Safety Council, *Forging Safety Manual*, 1990; E. G. Nisbett, *Steel Forging*, STP903, 1986.

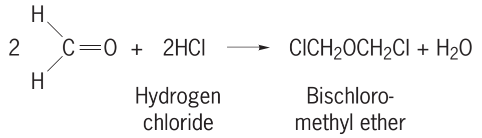
Formaldehyde

The simplest aldehyde, formula HCH═O. Because of its extreme reactivity, even with itself, it cannot be readily isolated or handled in the pure state. Therefore, it is produced and marketed as an aqueous solution (usually 37–50% formaldehyde by weight), sometimes known as Formalin. It is also sold as the solid hydrated polymer known as paraformaldehyde or paraform. Approximately 60% of all the formaldehyde produced is for captive use. *See* ALDEHYDE.

Properties. Formaldehyde undergoes many of the general reactions which are typical of aldehydes. Some interesting reactions which are somewhat specific for formaldehyde are (1) and (2). Bischloromethyl ether has been classified by the U.S. Occupational Safety and Health Administration as a carcinogen.

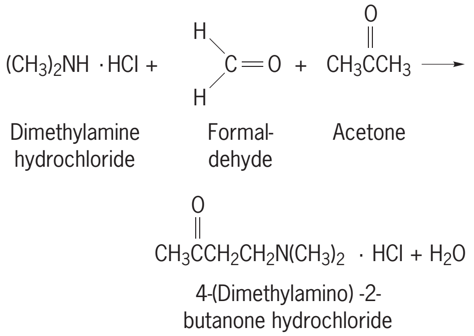


(1)

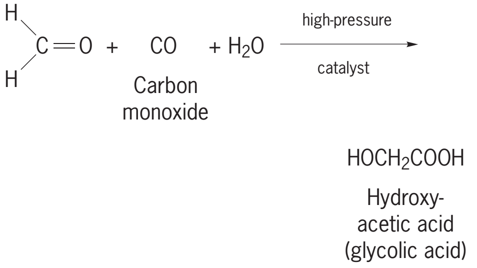


(2)

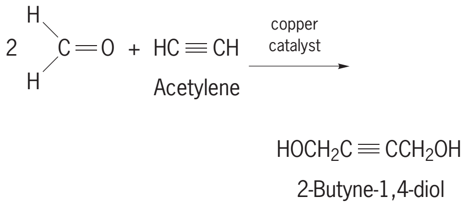
The Mannich reaction (3) involves the reaction of an amine hydrochloride with formaldehyde and a ketone. Reaction of formaldeyde with carbon monoxide is shown in reaction (4). Reaction (5) represents the Reppe ethynylation reaction.



(3)



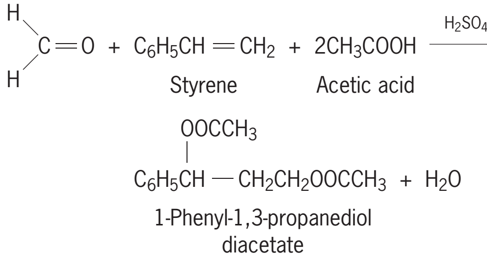
(4)



(5)

The Prins reaction (6) involves the reaction of formaldehyde with olefinic aromatic hydrocarbons in the presence of an organic acid and a mineral acid catalyst.

The reaction of formaldehyde with phenols, urea, and melamine to form resins proceeds first with



$$\mathrm{HCH{=}O} + C_6H_5CH{=}CH_2 + 2CH_3COOH \xrightarrow{H_2SO_4} C_6H_5\underset{}{\overset{\displaystyle OOCCH_3}{\overset{|}{C}}}H-CH_2CH_2OOCCH_3 + H_2O \qquad (6)$$

Styrene Acetic acid 1-Phenyl-1,3-propanediol diacetate

formation of methylol (—CH_2OH) derivatives followed by intermolecular dehydration to produce methylene (CH_2) linkages. This reaction is catalyzed by either alkali or acid, and the extent of resinification is controlled by temperature, reaction time, and concentration. *See* PHENOLIC RESIN.

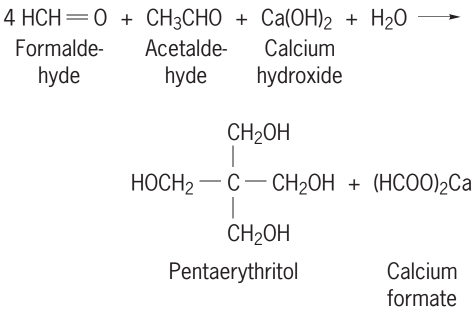
Production. Most of the formaldehyde is manufactured by the oxidation of methanol with air over a metal catalyst in the temperature range of 450–650°C (840–1200°F). Catalysts are either silver, copper, or an iron-molybdenum mixture. The reaction is shown as (7).

$$2CH_0H + O_2 \longrightarrow 2HCH{=}O + 2H_2O \qquad (7)$$

Methanol

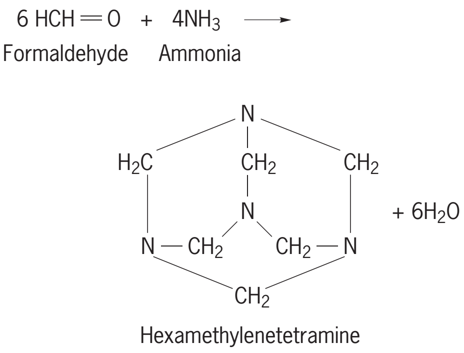
Increasing quantities of formaldehyde are produced by the partial oxidation of hydrocarbons from natural gas. Air is used as the source of oxygen, excess hydrocarbon to minimize complete oxidation, and steam to quench the reaction. The yields of formaldehyde are low, but the raw materials are plentiful and low in cost. Since other products, such as alcohols, acids, ketones, and other aldehydes, are formed, the purification of formaldehyde produced by this process may involve fractionation, liquid-liquid extraction, azeotropic distillation, and extractive distillation.

Uses. Formaldehyde is used principally to produce synthetic resins and adhesives by reaction with phenols, urea, and melamine. This use accounts for about 75% of the total production. Approximately 15% is used in the manufacture of textiles, dyes, drugs, paper, leather, photographic materials, embalming agents, disinfectants, and insecticides. It is also used for the industrial production of pentaerythritol and hexamethylenetetramine, reactions (8) and (9).



$$4\,HCH{=}O + CH_3CHO + Ca(OH)_2 + H_2O \longrightarrow HOCH_2-\overset{\displaystyle CH_2OH}{\underset{\displaystyle CH_2OH}{C}}-CH_2OH + (HCOO)_2Ca \qquad (8)$$

Formaldehyde Acetaldehyde Calcium hydroxide Pentaerythritol Calcium formate



(9)

L. Madestau

Formation

A fundamental geological unit used in the description and interpretation of layered sediments, sedimentary rocks, and extrusive igneous rocks. A formation is defined on the basis of lithic characteristics and position within a stratigraphic succession. It is usually tabular or sheetlike, and is mappable at the Earth's surface or traceable in the subsurface (for example, between boreholes or in mines). Examples are readily recognized in the walls of the Grand Canyon of northern Arizona (see **illus.**). Each formation is referred to a section or locality where it is well developed (a type section), and assigned an appropriate geographic name combined with the word formation or a descriptive lithic term such as limestone, sandstone, or shale (for example, Temple Butte Formation, Hermit Shale). This usage of "formation" by geologists differs from its informal lay usage for stalactites, stalagmites, and other mineral buildups in caves. *See* STRATIGRAPHY.

Distinctive lithic characteristics used to designate formations include chemical and mineralogical composition, particle size and other textural features, primary sedimentary or volcanic structures related to processes of accumulation, fossils or other organic content, and color. Lithology may be interpreted also in boreholes from cores and cuttings produced by drilling, and from electrical, radioactive, seismic, or other properties measured by instruments lowered into the holes. Contacts or boundaries between formations are chosen at surfaces of abrupt lithic change or within zones of gradational lithic character. Commonly, these contacts correspond with recognizable changes in topographic expression, related to variations in resistance to weathering. For example, massive carbonate rocks and sandstones, such as the Redwall Limestone and Coconino Sandstone in the Grand Canyon, tend to be cliff-forming, whereas shales such as the Bright Angel form ledges or areas of low topography because clay minerals are easily weathered. Formations normally possess some degree of internal lithic homogeneity, but they may be composed of repetitions of two or more different kinds of sediments or rocks, or extreme heterogeneity may itself constitute a distinctive feature.

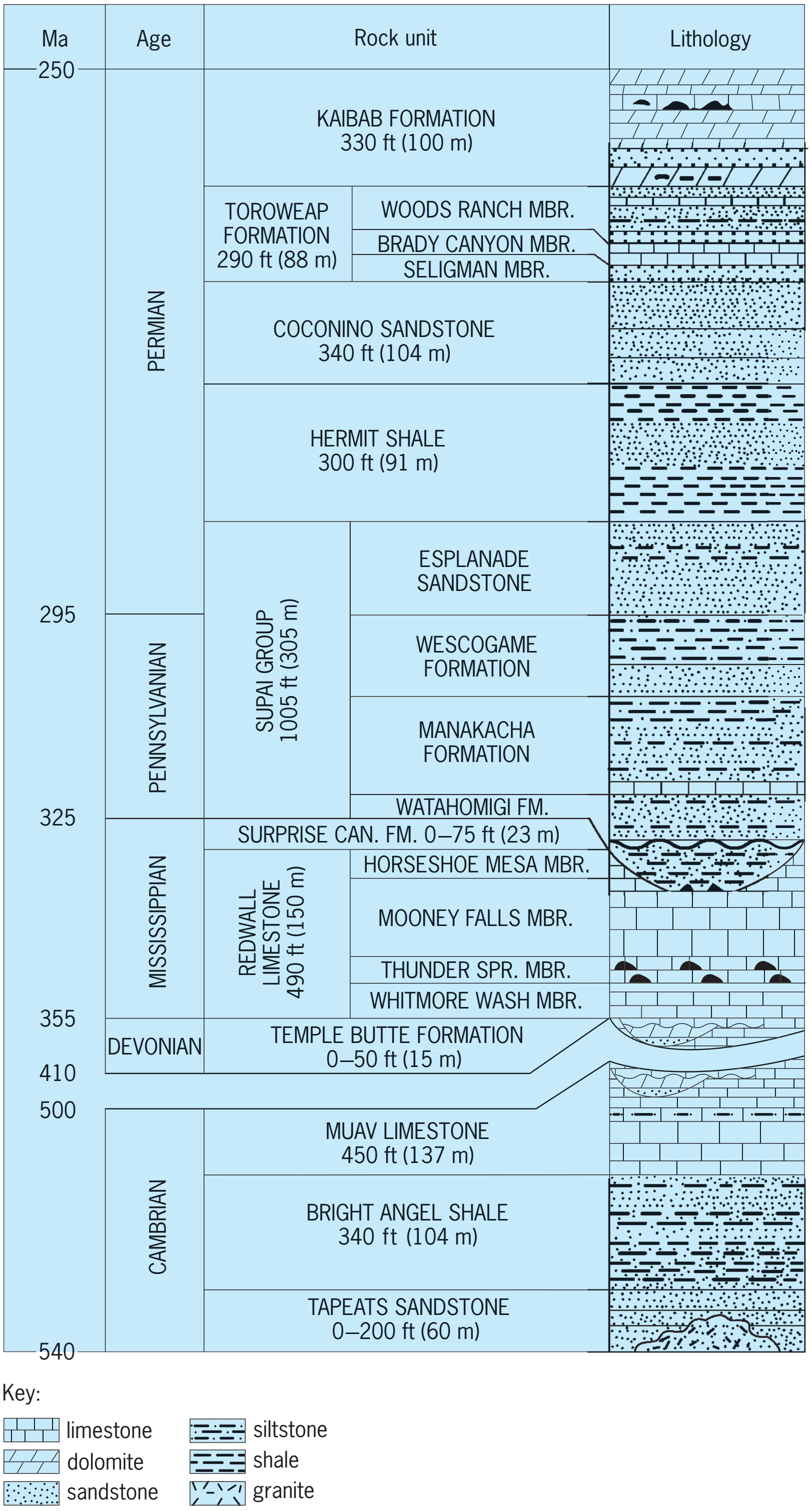
Different formations in a sedimentary succession may also be lithologically similar, but useful to distinguish because they are located at different stratigraphic levels. *See* SEDIMENTARY ROCKS; SEDIMENTOLOGY.

Formations tend to be tabular because most sedimentary layers are characterized by gradual lateral changes in lithic character. Exceptions are found at the seaward margins of ancient continental shelves and carbonate platforms, where abrupt lateral changes in paleowater depth are reflected by abrupt changes in lithology; and along the faulted margins of some sedimentary basins, where coarse-grained alluvial gravels or conglomerates pass laterally into comparatively fine-grained lacustrine or marine deposits.

Mappability is an essential characteristic of a formation because such units are used to delineate geological structure (faults and folds), and it is useful to be able to recognize individual formations in isolated outcrops or areas of poor exposure. Well-established formations are commonly divisible into two or more smaller-scale units termed members and beds (for example, subdivisions of the Redwall Limestone). In other cases, formations of similar lithic character or related genesis are combined into composite units termed groups and supergroups (for example, Supai Group). The rank of a named unit may vary from one area to another (for example, from group to formation) according to whether or not subunits are readily mappable. Changes in rank are also justified in the light of new geological knowledge.

Formations vary in thickness from a feather edge at a depositional or erosional limit to several kilometers. They tend to be thin where successions are characterized by abundant and easily recognized variations in lithology, and in areas that have been the focus of many decades or even centuries of geological investigations (for example, southern England). Formations tend to be thicker in successions that are characterized by limited lithic variability or that have been subjected only to reconnaissance geological mapping [for example, basinal shales, successions in remote areas, and strata of Precambrian age (older than 543 million years)].

Although commonly used as a framework for interpreting geological history, formations and related units are conceptually independent of geological time. They may represent either comparatively short or comparatively long intervals of time. Accumulation of a particular unit may have begun earlier in some places than in others, and the time span represented by a unit may be influenced by later erosion. In some cases, a formation cropping out at one locality may be entirely older or younger than the same lithic unit at another locality. Normal practice is to avoid defining formations to include regional erosion surfaces, but recent efforts to locate such discontinuities in sedimentary successions have inevitably led to the recognition of breaks in sedimentation within many of the formations defined in earlier studies.



Geological units of Paleozoic age recognized in the walls of the Grand Canyon, Arizona. (*Modified from S. S. Beus, Geology along the South Kaibab Trail, eastern Grand Canyon, Arizona, in S. S. Beus, ed., Geological Society of America Centennial Field Guide, vol. 2: Rocky Mountain Section, Geological Society of America, pp. 371–378, 1987*)

Although the concept of time plays no role in the definition of a formation, evidence of age is useful in the recognition of lithologically similar units far from their type localities.

Nicholas Christie-Blick

Bibliography. North American Commission on Stratigraphic Nomenclature, North American Stratigraphic Code, *Amer. Ass. Petrol. Geol. Bull.*, 67:841–875, 1987; A. Salvador (ed.), *International Stratigraphic Guide*, 2d ed., Geological Society of America, 1994.

Fossil

A record of earlier life buried in rock. Originally meaning any distinctive object that has been dug up (from Latin *fodio*, dig), the term "fossil" soon came to refer particularly to things resembling animals and plants. These objects were known before it was recognized that life on Earth has a long history. They were generally interpreted as having somehow grown in the rock or been placed there by a creator, or as being remains of organisms living somewhere else at the time in undiscovered seclusion.

In the eighteenth and nineteenth centuries, the realization grew that fossils indeed represent a record of older life, and the success of Darwin's theory of evolution provided the possibility for scientific investigation of life's history through paleontology. Fossils are now known from almost the whole time span of Earth's sedimentary record, and they provide a panoramic picture of how the biosphere has changed through time, gradually as well as in bursts of evolutionary activities and catastrophic extinctions.

Fossilization. A widespread conception is that a fossil is a shell or skeleton turned into stone. This picture is not false, but it represents only a few of the ways in which life can leave its trace in Earth's accreting skin of sedimentary rocks. Any part of an organism can get preserved, not only hard parts but also soft tissues, cells, and organelles. Even when nothing of the original organisms is preserved, impressions and traces in the sediment give important information about their former presence, activities, and ecological roles. Also, fossilization can imply everything from preservation of the almost unaltered original tissues to their complete replacement with sediment or minerals growing in place. Organic molecules can be preserved, though in a more or less degraded state.

The biosphere normally recycles all organic and inorganic matter produced by organisms; fossils represent dead individuals that to some degree escaped that process. Most decomposition is by aerobic scavengers, fungi, and bacteria, and so a prerequisite for fossilization is that the dead body is quickly and permanently subjected to an environment in which decomposers cannot be active. A combination of anoxic water and rapid sedimentation is a typical condition favorable to fossilization, though other conditions, such as extreme temperatures, salinity, poisonous environment, desiccation, or rapid mineralization, are also known to promote fossilization. Some kind of microbial activity, however, seems to be a prerequisite for many types of fossilization, particularly of soft tissue. *See* EDIACARAN BIOTA.

Fig. 1. Skull of sabre-toothed cat, *Amphimachairodus*, from the Miocene Epoch. (*Palaeontological Museum, Uppsala University*)

Shells and skeletons. Mineralized hard parts, such as shells, spicules, and bones, are by far the most common type of fossil (**Fig. 1**). Although they typically contain a substantial proportion of organic material, their mineral phase usually ensures that they are more resistant than soft tissues to biological decomposers. The most common skeletal minerals are opal (hydrated silica), apatite (calcium phosphates), and calcite and aragonite (calcium carbonates). Apatite, the common constituent of bone, is particularly stable under normal conditions; and it is often preserved crystallographically intact in fossils, though replacement and infilling of the other components of the hard tissue with minerals (permineralization) has usually taken place. Calcite is also often preserved with its original crystallography, whereas aragonite and opal tend to be recrystallized to calcite and quartz, respectively, or replaced altogether.

Even when none of the original hard tissue is preserved, its former presence may promote fossilization through its initial resistance to degradation. Shells are frequently preserved as molds or casts; for example, lithified infillings (internal molds) of mollusk shells are common fossils.

Some hard or protective tissues may not contain any appreciable mineral phase but are nevertheless resistant to degradation and may commonly be fossilized. Arthropod cuticle, cnidarian perisarc, hemichordate periderm, leaf cuticle, and spore exine are examples of such outer protective tissues. What gives them their resistance is usually tanned proteins, polysaccharides, or waxes.

Exceptional preservation. Hard or otherwise resistant tissues are thus responsible for the overwhelming majority of all fossils. Uncritical reliance on such skeletal fossils leads to grave biases with regard to the nature and anatomy of individual organisms and to the composition of the once-living biota. Only about one-tenth of the taxa living today have tissues and life habitats that make them readily fossilizable. For this reason, paleontology is increasingly dependent

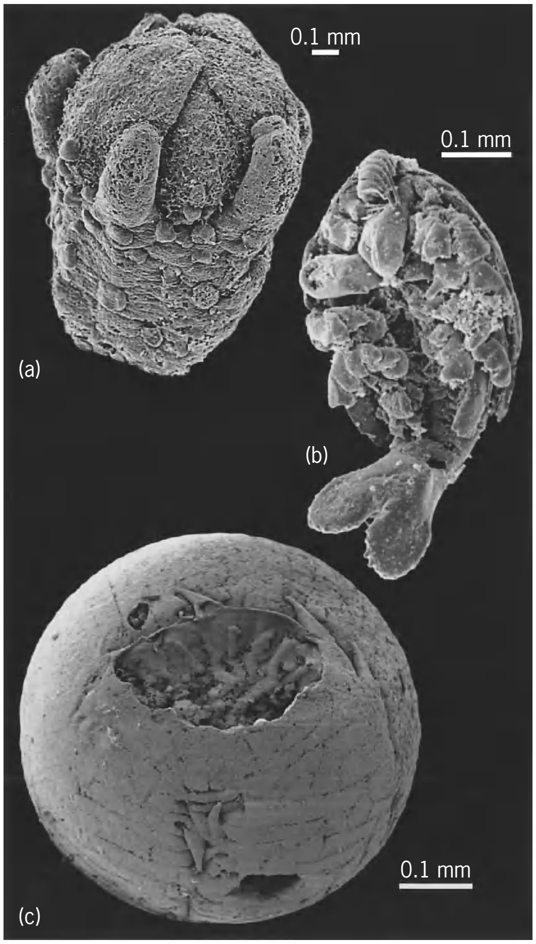


Fig. 2. Fossils with exceptional preservation of soft parts. (*a*) Flower from the Cretaceous Period, *Silvianthemum*, preserved by charring due to a forest fire. (*b*) Crustacean arthropod, *Rehbachiella*, and (*c*) embryo of segmented worm, *Markuelia*, both from the Cambrian Period and preserved by impregnation with calcium phosphate. (*Part a from E. M. Friis, Silvianthemum suecicum gen. et sp. nov., a new saxifragalaen flower from the Late Cretaceous of Sweden, Biologiske Skrifter, Det Kongelige Danske Videnskabernes Selskab, 36, plate 1, fig. 1, 1990; b from D. Walossek, The Upper Cambrian Rechbachiella and the phylogeny of Branchiopoda and Crustacea, Fossils and Strata, 32, plate 32, fig. 2, 1993; c reprinted with permission from S. Bengtson and Z. Yue, Fossilized metazoan embryos from the earliest Cambrian, Science, 277:1645 –1648, copyright 1997, American Association for the Advancement of Science*)

on sites with unusual conditions of preservation, allowing for the fossilization also of soft parts and of more complete samples of the total biota (**Fig. 2**).

Early silicification of sediments may trap and preserve biota; this is the main process responsible for knowledge of the microbially dominated biosphere of the Archean and Proterozoic eons, up to about 550 million years ago. The oldest known fossils on Earth are bacterialike structures from Archean sediments, dated to approximately 3.5 billion years. *See* PALEOECOLOGY; SEDIMENTOLOGY.

Another process known to promote exquisite preservation is impregnation with calcium phosphate during early diagenesis (physical and chemical changes occuring in sediments between deposition and solidification); this has been known to preserve soft tissues, even to cellular detail. Small arthropods have been preserved in profusion and in detail by phosphatization of the soft cuticular integument. Even early embryonic stages of animals are preserved through phosphatization; the record of animal embryos now goes back to the late Proterozoic.

Seemingly destructive forest fires may result in excellently preserved plant tissues through coalification. Wood fragments, seeds, and even delicate structures such as flowers can be thus preserved by charring.

Amber, fossilized tree resin, is well known for its capacity to trap and fossilize insects and other small animals and plant parts. Insects in amber are known from at least the last 100 million years.

Freezing has yielded spectacular finds of soft-tissue preservation of, for example, mammoths. The dependence on permanent low temperatures for maintaining the fossils, however, limits this kind of preservation to the most recent fossil biotas.

Various types of fine-grained shales and mudstones have for a long time been known to be potential sites of exquisitely preserved fossils. Examples are the Cambrian Burgess Shale (British Columbia) and Chengjiang mudstones (south China), the Devonian Hunsrück Shale, the Jurassic Holzmaden Shale [= Posidonienschiefer] and Solnhofen Lithographic Limestone (Germany), the Jurassic-Cretaceous Chaomidianzi shales of Liaoning (northeast China), and the Eocene Messel Oil Shale (Germany). These rocks are more or less compressed by the weight of overlying sediment, and so the fossils are not preserved in as full relief as in the other types of extraordinary preservation mentioned earlier. However, the shaley deposits are capable of preserving much larger fossils (such as the ichthyosaurs of Holzmaden and the feathered dinosaurs and birds of Liaoning) than most of the other processes. *See* BURGESS SHALE.

Other types of exceptional fossil preservation are known, though they are more incidental and may be restricted to a short stratigraphic interval. The late Proterozoic Ediacara biota of large, probably multicellular organisms is preserved in coarse-grained sediments spanning a relatively short interval of time; a suggested explanation of this pattern is that the preservation was dependent on the presence of ungrazed microbial mats on the sea floor and that these disappeared when grazing and burrowing animals became common.

Trace fossils (marks of animal activities in sediment) and coprolites (fossilized feces) generally give less information than body fossils about the anatomy of the ancient organisms, but they are important sources of ecological and behavioral information (**Fig. 3**). *See* TRACE FOSSILS.

Time indicators. Fossils were used to order and date rocks even before the principle of evolution was generally accepted, and throughout time the most important practical use of fossils has been biostratigraphy (the identification, ordering, and correlation of fossil-containing sedimentary rocks). Rocks from each slice of time represent a unique but diverse biosphere, and the preservation of the organisms is subject to complex interacting factors; therefore

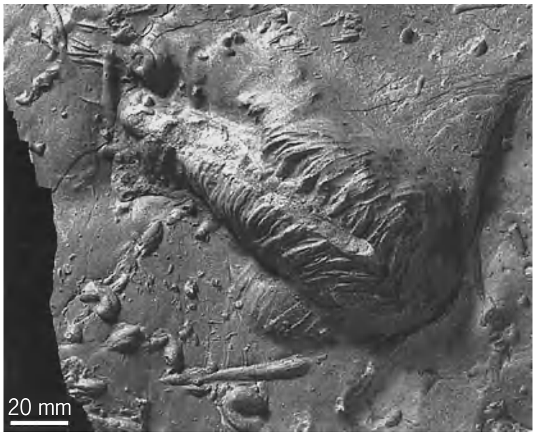


Fig. 3. Trace fossil, *Cruziana*, formed by an arthropod (probably a trilobite) digging for a wormlike animal in soft mud during the Cambrian Period. The picture shows the lower side of a sandstone bed casting of the original markings in the mud (now vanished). (*Swedish Geological Survey; from S. Jensen, Trace fossils from the Lower Cambrian Mickwitzia sandstone, south-central Sweden, Fossils and Strata, vol. 42, 1997*)

biostratigraphy combines analyses of systematics, evolution, ecology, sedimentology, and biogeography.

Biostratigraphy is further becoming integrated with other methods of correlating rocks, for example isotope geology, geomagnetism, and sequence stratigraphy. A central question in modern (bio) stratigraphy concerns what time resolution is possible, how much confidence can there be in known stratigraphical ranges of certain fossils, and how large are the errors in the dating of biological and geological events. *See* GEOLOGIC TIME SCALE; GEOLOGY.

Environmental indicators. Every organism has its range of tolerable environmental conditions, and so a key use for fossils is as indicators of parameters such as temperature, salinity, light intensity, and nutrition. Such paleoecological inferences are most reliable in the younger parts of the geological column, where there are organisms closely related to those living today, with known ecological preferences. In older rocks, inferences must be based more on analogs, functional morphology, and the character of the entombing sediment. Trace fossils are particularly important as environmental indicators, because they record activities that are tightly bound to the particular spot studied (no possibility of postmortem transport).

The chemistry of fossils is frequently used as environmental indicators; their isotopic composition can be used to infer, for example, sea-water temperature ($^{18}O/^{16}O$ ratios in shells) and diet ($^{13}C/^{12}C$ and $^{15}N/^{14}N$ ratios in bone); and trace elements in shells that grew by accretion may provide a proxy for changes in environmental chemistry during the life of the animal. *See* PALEOCLIMATOLOGY.

Global importance. The process of fossilization has a wider significance than just providing a record of past life. A considerable part of Earth's sedimentary cover is of biological origin, that is, fossils; and the permanent or temporary withdrawal of, for example, carbonate and organic matter from biogeochemical cycling by burial in sediment has a direct and profound influence on the composition of atmosphere and oceans and thus on the conditions for life. Stefan Bengtson

Fossil record. The fossil record is the sequence of life preserved as fossils in rocks of all ages in all parts of the Earth. The term is also used in a more restricted sense for the preserved history of a particular taxonomic group, as in "the fossil record of vertebrates"; the preserved history from a particular interval of time, as in "the Precambrian fossil record"; or the preserved history of a particular ecosystem, as in "the fossil record of hydrothermal vent communities."

Quality. The fossil record is vast, extending over 3.5 billion years and preserving representatives of all living kingdoms in all major ecosystems. But the record is also incomplete and knowledge of it is imperfect, despite more than two centuries of scientific investigation. Approximately 250,000 named animal species have been collected from the marine fossil record. The percentage that this represents of the total number of marine animals that ever existed can be estimated from the average duration of species and the approximate number of species extant at any given time. Average durations vary with taxonomic group, but 4 million years is probably a good average. The first animal appeared about 600 million years ago, and today there are about 250,000 species in the oceans. The path of increase in species numbers, or diversity, has been irregular, but can be roughly approximated as a straight line for the sake of calculation (**Fig. 4**). Given this assumption, the total number of marine animal species that ever lived is $^1/_2 \times 2.5 \times 10^5$ species $\times$ (600×10^6 years $\div$ 4×10^6 years) $= 188 \times 10^5$ species, and the percentage that have been described as fossils is 100 ×

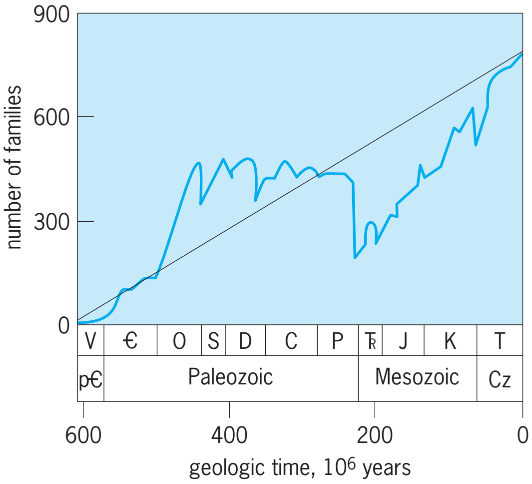


Fig. 4. Diversity of taxonomic families of animals in the marine fossil record, showing the irregular path of increase from the Vendian (that is, latest Precambrian) to the present. The increase of species diversity is thought to follow a similar path. The straight line shows a rough approximation of this diversity increase, with excess diversity after evolutionary radiations canceled by reduced diversity after mass extinctions. Geologic eras and periods are Precambrian (p Є), Cenozoic (Cz), Vendian (V), Cambrian (Є), Ordovician (O), Silurian (S), Devonian (D), Carboniferous (C), Permian (P), Triassic (Є), Jurassic (J), Cretaceous (K), and Tertiary (T). (*After N. Eldredge, ed. Systematics, Ecology, and the Biodiversity Crisis, Columbia University Press, 1992*)

$(2.5 \times 10^5$ species $\div 188 \times 10^5$ species$) = 1\%$. Thus, the known sample of the fossil record is small and will not permit a detailed reconstruction of the phylogeny of all life. However, for statistical and many other purposes, this sample is large, especially in comparison to political and consumer surveys.

Unfortunately, the sample from the fossil record is not random and unbiased. The quality varies greatly, from nearly complete sequences of populations of certain Miocene lake fish, to scattered, poorly preserved fossils of soft-bodied worms, to absence of recognized fossils for many bacterial groups. The fossil record is best for invertebrate animals with robust skeletons composed of calcium carbonate (for example, corals, clams, snails, and brachiopods) and poorest for small organisms without skeletons or durable cuticles (for example, many bacteria, protists, and fungi and various wormlike animals such as flat worms and nematodes). Between these extremes there is a wide spectrum of preservation that varies with the type of organism and its geologic age, environment, and biogeography.

In general, the fossil record is better known for younger eras because more sedimentary rock has survived the ravages of erosion, subduction, and metamorphism and because there has been more economic incentive for study. The vast majority of fossils have been obtained from Phanerozoic rocks, 0–545 million years in age. Archean rocks, 2.5–3.8 billion years in age, have so far yielded fewer than 10 deposits with verifiable fossil.

The Phanerozoic fossil record is more complete for marine environments, which are sites of net deposition, than for continental environments, which are sites of net erosion. Within the marine realm, the fossil record tends to be better for the continental shelf than for the deep sea, where most sediments older than 150 million years have been subducted or metamorphosed. On land, the record is much better for lowland areas, especially floodplains of coastal rivers, than for upland areas such as high plains, hills, and mountains where sedimentary deposits rarely are shielded from erosion for any significant length of time.

Even with preservation, sampling of the fossil record is quite variable. Organisms with multielement skeletons that dissociate after death, such as many echinoderms, are plentiful in the fossil record but are often very difficult to study (exceptions include many mammals and plants with easily identifiable teeth and pollen or spores, respectively). Preservable organisms that were rare, geographically restricted, or of short geologic duration tend to be sampled less frequently than abundant, widespread organisms; and organisms useful for stratigraphic correlation, such as foraminifers, ammonites, graptolites, conodonts, and pollen-producing plants, have been more intensively studied than organisms with limited utility, such as cyanobacteria, calcareous algae, and sponges. Finally, sampling of the fossil record has varied with political geography, with more intensive study in the industrialized world than in developing nations.

Overview of Phanerozoic record. Species are the basic units of evolution and the subject of study in detailed paleontologic investigations. However, most generalizations about the fossil record are made at higher taxonomic levels, particularly the genus and family. This is because complete data are easier to compile for the less numerous genera and families (approximately 40,000 and 4000, respectively, in the Phanerozoic marine record) and because sampling of higher taxa is more complete (only one species is needed to document the presence of a genus, family, and such). Figure 4 illustrates the record of diversity of animal families in the oceans and shows that in general the fossil record has become richer toward the Recent. **Figures 5** and **6** display summaries of family diversities within taxonomic classes for most marine and continental organisms. Diversities in these two figures are plotted as spindle diagrams that show the number of families through time graphed symmetrically about a central axis. These spindle diagrams show that the fossil record of taxonomic classes is highly skewed in terms of size, duration, and time of origin. Among marine animals (Fig. 5), most classes are small, and more than half of the recognized families are contained in only seven major classes, the Cephalopoda, Articulata (Brachiopoda), Crinoidea, Osteichthyes, Gastropoda, Bivalvia, and Trilobita (listed in order of decreasing size).

Approximately 60% of marine classes appear in the first quarter of the illustrated time (Vendian and Cambrian), whereas only 5% first appear in the last half (Permian to Pleistocene). Many of the early classes remain small and become extinct before the end of the Paleozoic Era. In contrast, those that radiate and attain substantial diversity during the Paleozoic tend to persist to the present day (the Trilobita being the most notable exception). However, there are many classes that have long histories but few fossil families. Most of these comprise soft-bodied animals (for example, Nemertina, Priapulida, Sipunculida) that are known from only a few extraordinary fossil deposits, termed *Lagerstätten*. In the marine record, 20% of fossil classes are known from only three *Lagerstätten*: the Cambrian Burgess Shale, the Devonian Hunsrück Shale, and the Carboniferous Francis Creek Shale of the Mazon Creek area.

Classes in the continental fossil record (Fig. 6) are similarly skewed. More than half of the 39 animal classes contain fewer than 7 fossil families, and more than 40% of the 2500 recognized families belong to a single class, the Insecta. From the Carboniferous to present, this class contains more families than all others in the continental fossil record, including the sum of all vertebrate classes. As in the marine record, exceptional views of the diversity of continental life are provided by a few *Lagerstätten*, notably the Oligocene Baltic ambers which have yielded an unprecedented number of nonmarine arthropods.

The continental fossil record of animals begins in the Silurian, later than the marine record, which begins in the Vendian. As in the marine record, first appearances of classes tend to be early, with two-thirds appearing before the end of the Carboniferous. The record for terrestrial plants (Fig. 6, bottom) is similar: most classes appear early (except

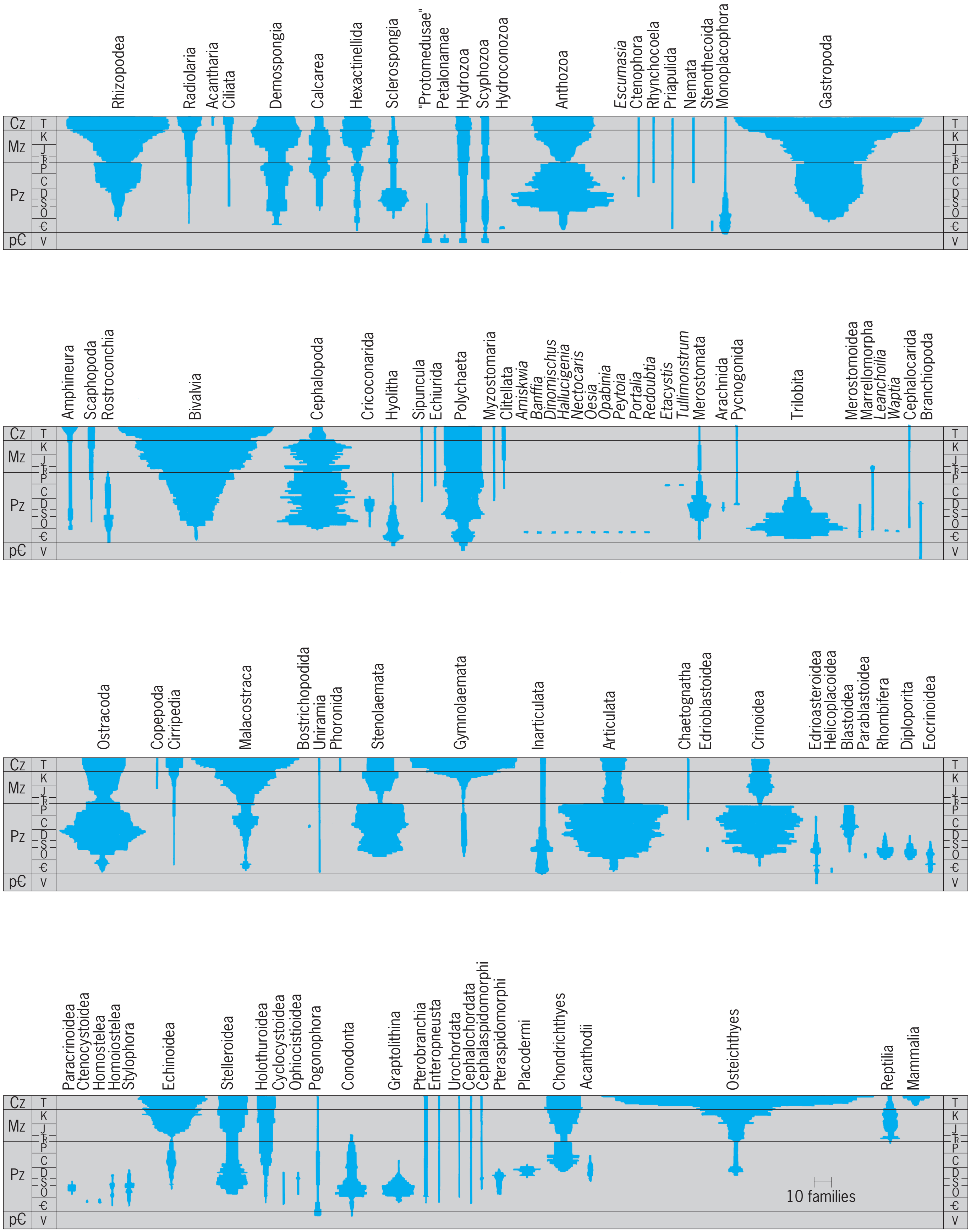


Fig. 5. Spindle diagrams displaying the diversity of families within animal classes in the marine fossil record. Time runs vertically in each strip from Vendian to Pleistocene. (Modern families without fossil representatives are not included.) Geologic eras and systems are Precambrian (p Є), Paleozoic (Pz), Mesozoic (Mz), Cenozoic (Cz). (*After J. W. Valentine, ed., Phanerozoic Diversity Patterns: Profiles in Macroevolution, Princeton University Press, 1985*)

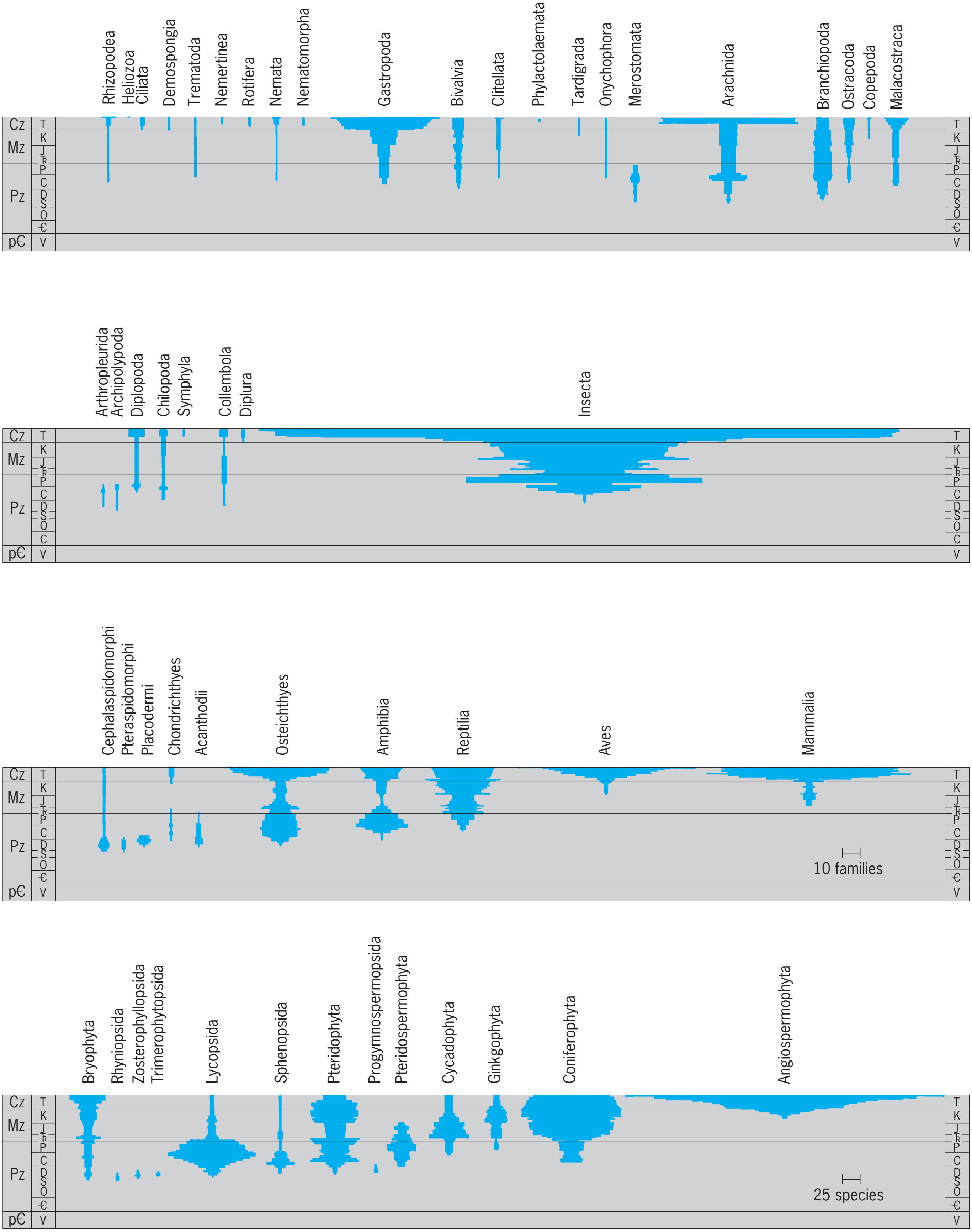


Fig. 6. Spindle diagrams displaying the diversity of families within animal classes and of species within plant classes in the continental (that is, nonmarine) fossil record. Plotting conventions are the same as in Fig. 5. (*After J. W. Valentine, ed., Phanerozoic Diversity Patterns: Profiles in Macroevolution, 1985*)

the Magnoliophyta = angiosperms), and only a few classes are large (notably the Pinophyta and Magnoliophyta). *See* PALEOBOTANY; PALEONTOLOGY.

J. John Sepkoski, Jr.

Bibliography. R. C. Moore et al. (eds.), *Treatise on Invertebrate Paleontology*, 1953–1998; D. R. Prothero, *Bringing Fossils to Life: An Introduction to Paleobiology*, 1998; D. M. Raup and S. M. Stanley, *Principles of Paleontology*, 2d ed., 1978; G. G. Simpson, *The Major Features of Evolution*, 1953; S. M. Stanley (ed.), *Macroevolution: Pattern and Process*, 1979; J. W. Valentine (ed.), *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, 1985.

Fossil apes

Apes and humans are closely related primates in the superfamily Hominoidea. The living hominoids are subdivided into the families Hylobatidae and Hominidae. The hylobatids or lesser apes (genus *Hylobates*) are represented by approximately nine species found throughout Southeast Asia. Humans and the great apes—the orangutan (*Pongo pygmaeus*), the gorilla (*Gorilla gorilla*), the common chimpanzee (*Pan troglodytes*), and the pygmy chimpanzee or bonobo (*Pan paniscus*)—are grouped in the Hominidae. In the past, the great apes were included in a separate family, the Pongidae, but subsequent anatomical and molecular studies showed that the African apes (*Gorilla* and *Pan*) are more closely related to humans than they are to the Asian orangutan.

The classification of fossil and living apes, along with their geological ages and geographical regions, follows. The asterisk indicates extinct apes.

- Superfamily Hominoidea
  - Family Hylobatidae
    - *Hylobates* (Pleistocene to Recent, Southeast Asia)
  - Family Hominidae
    - Subfamily Kenyapithecinae
      - *Griphopithecus** (middle Miocene, Turkey and Austria)
      - *Kenyapithecus** (middle Miocene, East Africa)
    - Subfamily Dryopithecinae
      - *Dryopithecus** (middle to late Miocene, Europe)
      - *Oreopithecus** (late Miocene, Italy)
    - Subfamily Ponginae
      - *Ankarapithecus** (late Miocene, Turkey)
      - *Gigantopithecus** (late Miocene to Pleistocene, Asia)
      - *Lufengpithecus** (late Miocene, China)
      - *Pongo* (Pleistocene to Recent, Southeast Asia)
      - *Sivapithecus** (middle to late Miocene, Indo-Pakistan)
    - Subfamily Homininae
      - *Ardipithecus** (early Pliocene, Ethiopia)
      - *Australopithecus** (Pliocene, Africa)
      - *Gorilla* (Recent, Africa)
      - *Graecopithecus* (= *Ouranopithecus*)* (late Miocene, Greece)
      - *Homo* (late Pliocene to Recent, cosmopolitan)
      - *Pan* (Recent, Africa)
      - *Paranthropus** (Pliocene to Pleistocene, Africa)
      - *Samburupithecus** (late Miocene, Kenya)
    - Uncertain family status
      - *Morotopithecus** (early Miocene, Uganda)
      - *Otavipithecus** (middle Miocene, Namibia)
- Uncertain superfamily status
  - Family Proconsulidae
    - *Afropithecus** (early Miocene, Kenya)
    - *Dendropithecus** (early Miocene, East Africa)
    - *Heliopithecus** (early Miocene, Saudi Arabia)
    - *Kalepithecus** (early Miocene, East Africa)
    - *Kamoyapithecus** (late Oligocene, Kenya)
    - *Limnopithecus** (early to middle Miocene, East Africa)
    - *Mabokopithecus** (middle Miocene, Kenya)
    - *Micropithecus** (early and middle Miocene, East Africa)
    - *Nyanzapithecus** (early and middle Miocene, East Africa)
    - *Proconsul** (early and middle Miocene, East Africa)
    - *Rangwapithecus** (early Miocene, East Africa)
    - *Simiolus** (early Miocene, Kenya)
    - *Turkanapithecus** (early Miocene, Kenya)

Evolutionary history. The evolutionary history of the extant hominoids is poorly known, with the notable exception of humans, which have a relatively complete fossil record extending back more than 4 million years. The earliest fossil apes that can be definitively linked to the modern hylobatids are known from sites in China dated to less than 1.5 million years ago (Ma), while the fossil record for the African apes is entirely unknown. The evolution of the orangutan is, by comparison, much better documented. Fossil teeth from cave sites in Asia dating back more than 1 million years show that orangutans in the past were considerably larger than they are today, and that unlike their living relatives, which are found only on Sumatra and Borneo, they once had a wider distribution in Southeast Asia that extended as far north as China. In contrast to the paucity of fossils available to trace the evolutionary history of hominoids over the past 5 million years, there is a wealth of evidence from the Miocene Period (23–5 Ma). This evidence shows that apes were once much more common and more diverse than they are today.

Proconsulids. The remains of apelike fossil primates, commonly known as proconsulids, have been recovered from sites in Kenya, Uganda, and Saudi Arabia dating to the early Miocene (23–16 Ma). There are 15 species of proconsulids, ranging in size from the tiny *Micropithecus clarki* (4 kg; 9 lb) to the chimpanzee-sized *Proconsul major* (50 kg; 110 lb). Comparisons of their teeth, jaws, and skeletons indicate that proconsulids exhibited a wide diversity of dietary and locomotor behaviors, but they were typically arboreal and ate various combinations of soft fruits and young leaves. The best-known species are

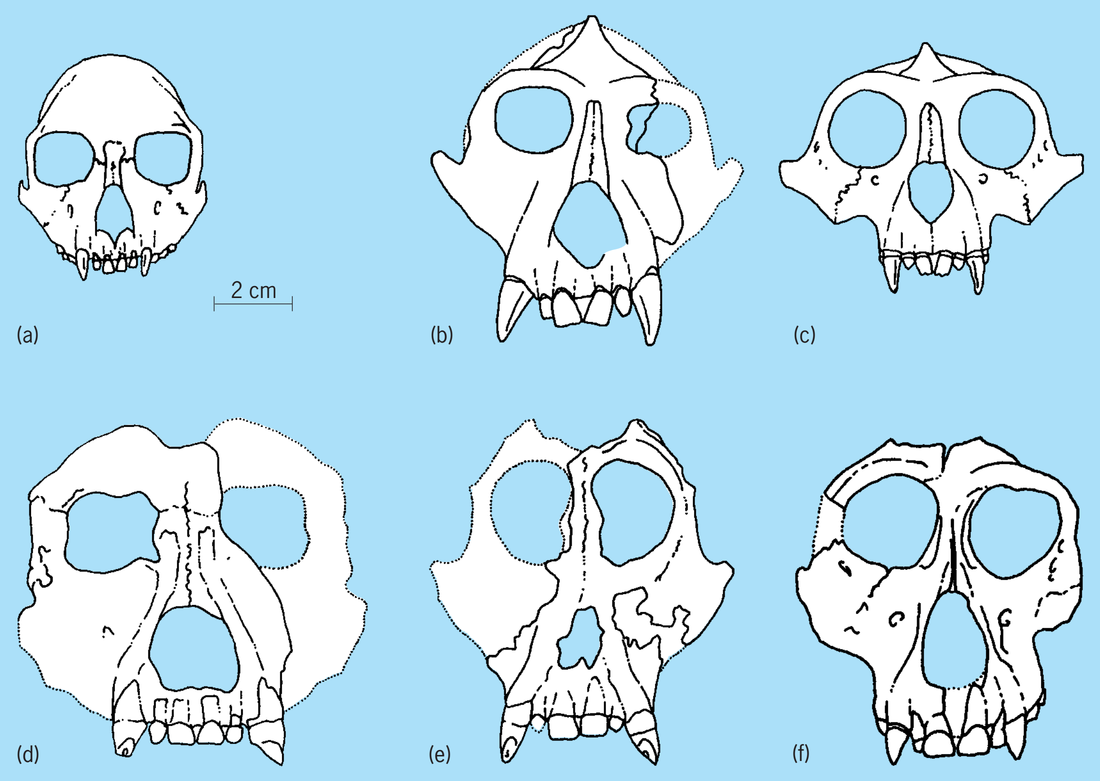


Fig. 1. **Reconstructions and partial reconstructions of the skulls of fossil apes: (*a*) *Proconsul*, (*b*) *Afropithecus*, (*c*) *Oreopithecus*, (*d*) *Graecopithecus*, (*e*) *Sivapithecus*, and (*f*) *Ankarapithecus*.**

Proconsul heseloni, *Turkanapithecus kalakolensis*, and *Afropithecus turkanensis* (**Fig. 1**). Studies have shown that proconsulids represent either the earliest known hominoids or primitive stem catarrhines (the group which gave rise to both Old World monkeys and apes). They are certainly more primitive than any of the living apes, retaining generalized skulls and teeth, and monkeylike postcranial skeletons. However, during the early Miocene there was at least one species of hominoid living in East Africa, *Morotopithecus bishopi*, which had already acquired some of the unique features of modern apes. This species, from Moroto in Uganda (dated to more than 20 Ma), had specializations of the lumbar vertebrae and scapula, not found in the contemporary proconsulids, that indicate that it had developed the stiff-backed, partially upright posture and suspensory forelimbs that are characteristic of modern hominoids.

During the middle Miocene (16–10 Ma), conditions in East Africa became somewhat drier, cooler, and more seasonal, and open woodland habitats replaced the humid tropical forests that were typical of the early Miocene. These ecological changes coincided with the appearance in East Africa of a more advanced type of hominoid, *Kenyapithecus*. This ape is derived relative to the proconsulids in having thickened enamel on its cheek teeth, more robust jaws, and relatively larger upper premolars. The limb bones indicate that *Kenyapithecus* was more terrestrially adapted than proconsulids and exhibited a number of specialized features that link it more closely to extant hominoids. Another intriguing fossil ape from this time period is *Otavipithecus* from Namibia, dated to approximately 13 Ma, and the only Miocene hominoid recorded from southern Africa. Until recently, it was known only by a single jaw fragment, so its relationships to other fossil and extant hominoids have been difficult to establish. However, additional cranial and postcranial specimens of *Otavipithecus* have now been discovered, and it is hoped that these will shed further light on its affinities.

Eurasian hominoids. Until the middle Miocene, hominoids were restricted to Africa, but during this period they migrated into Eurasia. The earliest Eurasian hominoid, dated to 16–15 Ma, is *Griphopithecus*, which is known from Turkey and central Europe. The teeth and jaws are similar to those of *Kenyapithecus* from East Africa, to which it is probably closely related. Once in Eurasia, hominoids became established over a wide geographical region, extending from Spain in western Europe to eastern China, and they became increasingly diversified during the middle and late Miocene (16–5 Ma). The best-known fossil Eurasian hominoids are *Dryopithecus* (western and central Europe), *Oreopithecus* (Italy), *Graecopithecus* or *Ouranopithecus* (Greece), *Ankarapithecus* (Turkey), *Sivapithecus* (Indo-Pakistan), and *Lufengpithecus* (China) [Fig. 1]. Of these forms, *Sivapithecus* is evidently closely related to the living orangutan, but the relationships of the other Eurasian Miocene hominoids remain contentious. A number of alternative hypotheses about the interrelationships of Eurasian Miocene hominoids have been proposed: (1) that they form

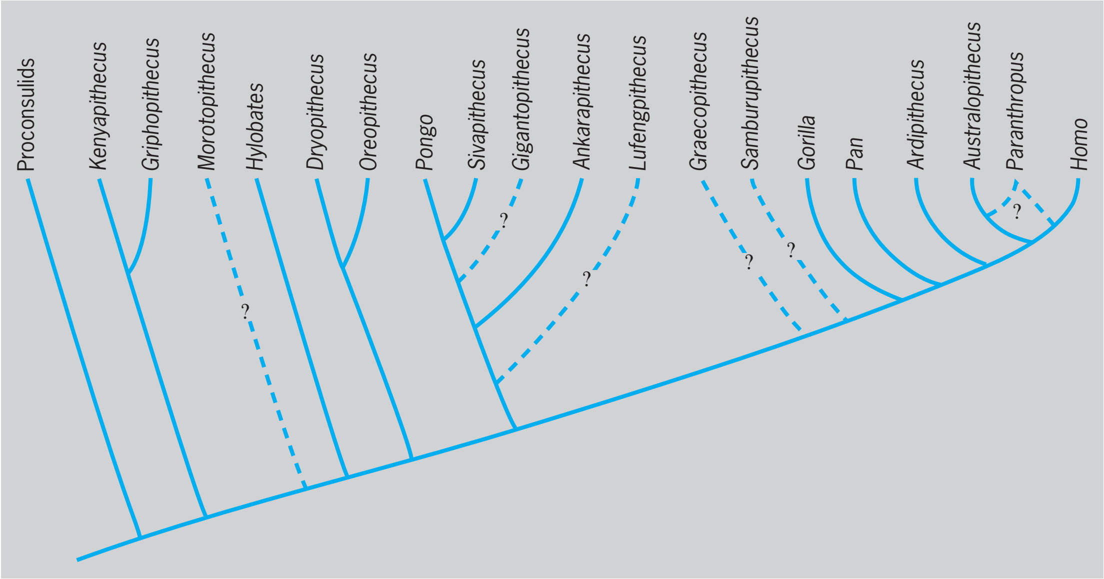


Fig. 2. Cladogram of the evolutionary relationships of fossil and living apes. Broken lines indicate uncertain relationships.

a closely related group with the living orangutan, all being derived from a common ancestor that migrated into Europe from Africa sometime during the middle Miocene; (2) that some of the Eurasian hominoids, such as *Dryopithecus* and *Graecopithecus*, are more closely related to the African apes and humans than they are to *Sivapithecus* and the orangutan; and (3) that they represent a diverse group containing primitive hominoids, as well as forms belonging to the orangutan and African great ape lineages (**Fig. 2**).

An ecological shift from moist temperate woodlands to drier, more seasonal habitats during the later Miocene coincided with a sharp decline in the diversity of hominoids in Eurasia. The only survivor in Europe toward the end of the Miocene was *Oreopithecus*, a highly specialized relative of *Dryopithecus* that was isolated on a group of islands in the northern Mediterranean which today form part of Italy (Fig. 1). *Lufengpithecus* and *Sivapithecus*, along with *Gigantopithecus*, are found in the late Miocene of Asia. The ape *Gigantopithecus* was the largest known hominoid, with massive jaws and teeth specialized for eating tough, fibrous vegetation, and with an estimated body weight that may have exceeded 200 kg (440 lb) [living male gorillas average only 170 kg or 375 lb]. All of these Eurasian hominoids became extinct by the close of the Miocene, except for *Gigantopithecus*, whose remains have been recovered from Pleistocene cave sites in China dated to less than 1 Ma.

Hominoids also became extremely rare in Africa during the late Miocene. A large hominoid, *Samburupithecus*, known only by a single maxilla from Kenya (dated to 10–8 Ma), may represent a close relative of the African apes and humans. A few isolated teeth and a lower jaw fragment of fossil hominoids from the late Miocene sites of Lukeino and Lothagam (dated to 7–5 Ma) in northern Kenya have been suggested to be the earliest known occurrence of the human lineage, but the remains are too scrappy to be certain of their taxonomic affinities. The earliest definitive record of fossil hominoids that are more closely related to humans than they are to the African great apes is known from the Pliocene (5.2–1.6 Ma) with the appearance of *Ardipithecus ramidus* from Ethiopia (4.4 Ma), *Australopithecus anamensis* from Kenya (4.2–3.9 Ma), and *Australopithecus afarensis* from Ethiopia and Tanzania (4–3 Ma). *See* APES; FOSSIL HUMANS; MAMMALIA; MONKEY.

Terry Harrison

Bibliography. D. R. Begun, C. V. Ward, and M. D. Rose, *Function, Phylogeny, and Fossils: Miocene Hominoid Evolution and Adaptations*, 1997; R. L. Ciochon and R. S. Corruccini, *New Interpretations of Ape and Human Ancestry*, 1983; J. G. Fleagle, *Primate Adaptation and Evolution*, 2d ed., 1998; R. G. Klein, *The Human Career*, 1989; B. Wood, L. B. Martin, and P. Andrews, *Major Topics in Primate and Human Evolution*, 1986.

Fossil fuel

Any naturally occurring carbon-containing material which when burned with air (or oxygen) produces (directly) heat or (indirectly) energy. Fossil fuels can be classified according to their respective forms at ambient conditions. Thus, there are solid fuels (coals); liquid fuels (petroleum, heavy oils, bitumens); and gaseous fuels (natural gas, which is usually a mixture of methane, CH_4, with lesser amounts of ethane, C_2H_6, hydrogen sulfide, H_2S, and numerous other constituents in small proportions).

Some differentiation must be made between the natural fossil fuels and those synthetic fuels derived by incurring major changes in the character of the fossil fuel. Examples are the liquid products made from coal by any one of a variety of potential process options; such materials are synthetic liquid fuels.

Heating values of representative fuels

Fossil fuel	Btu/lb	Heating value		
		Btu/ft^3	MJ/k^3	MJ/m^3
Natural gas		900		33.5
Petroleum	19,000		44.1	
Heavy oil	18,000		41.8	
Tar-sand bitumen	17,800		41.3	
Coal				
Lignite	8,000*		18.6	
Subbituminous	10,500*		24.4	
Bituminous	15,500*		36.0	
Anthracite	15,000*		34.8	

*Representative values are given because of the spread of subgroups with various heating values.

See COAL GASIFICATION; COAL LIQUEFACTION; SYNTHETIC FUEL.

Occurrence. Coal is a naturally occurring organic rock that originated from material which accumulated from the prolonged (millions of years) decay of woody plants. Coal origins are typically represented as beginning in a peat swamp in which trees, ferns, and the like are deposited, and buried by sand, silt, and mud. As a result of temperature and pressure effects, metamorphosis of the woody material occurs to produce the various types of coal. A large proportion of existing coal originated from materials deposited during the Carboniferous Period some 2.8–3.5 × 10^8 years ago. *See* COAL.

Petroleum occurs in highly porous rock reservoirs that are enclosed by an impermeable basement rock and an impermeable cap rock. The petroleum originated from the deposition of dead microscopic plant and animal organisms in ancient marine sediments. Geological effects such as temperature, time, and the pressure of overburden brought about chemical changes resulting in the production of hydrocarbon materials. In addition, it is generally recognized that the reservoir rock and the source rock are not the same. Migration occurred, for example, because of overburden pressures, from the source rock (now usually shales and clays) into the more coarse-grained reservoir rock (usually sandstones or carbonate rocks). The source rocks range in age up to about 5 × 10^8 years, but it is believed that conversion of the source material occurred during a period of 1–3 × 10^6 years. *See* PETROLEUM; PETROLEUM GEOLOGY.

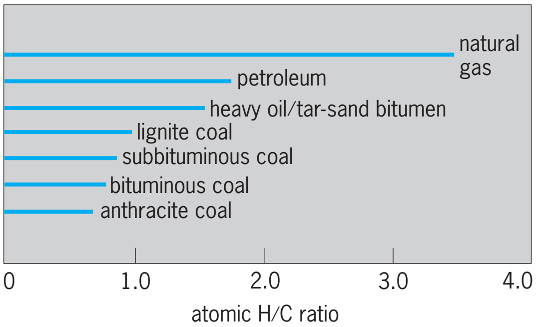
Heavy oils and tar-sand bitumens (natural asphalts) are a poorer-quality petroleum that has come into prominence because of the dwindling supplies of conventional petroleum. Both had origins similar to that of conventional petroleum but are believed to be younger (less mature), and occur in sandstone or carbonate reservoirs. *See* ASPHALT AND ASPHALTITE; BITUMEN; OIL SAND; OIL SHALE.

Like conventional petroleum, heavy oils occur in fields, but unlike conventional petroleum, heavy oils usually require additional (thermal) stimulation for recovery at the surface. Tar-sand bitumens occur as deposits in which the ore body (reservoir rock) may contain 1–20% by weight bitumen. The bitumen is immobile at the formation conditions and occurs as a thick, black, near-solid natural asphalt. The bitumen can be recovered by mining followed by removal from the associated sand (hot-water separation process). Nonmining (in-place) techniques require severe thermal or chemical treatments to stimulate bitumen mobility in the reservoir and recovery at the surface.

Like petroleum, natural gas also occurs in subterranean porous reservoir rocks either alone or in association with petroleum. The traps also consist of impermeable basement and cap rocks. Natural gas also probably resulted from the decomposition of dead microscopic organisms which had been deposited in aqueous environments. It may also be produced as an integral part (or by-product) of the maturation process in which source material is converted to petroleum. *See* NATURAL GAS.

Composition. Fossil fuels are composed of carbon (C) and hydrogen (H) with varying amounts of nitrogen (N), oxygen (O), and sulfur (S). On an atomic basis, natural gas has the highest atomic H/C ratio and coal the lowest; petroleum has an atomic H/C ratio between the two. There are also varying H/C ratios within the coal and petroleum series. The atomic H/C ratios for various fossil fuels are compared in the **illustration**.

Coals vary in elemental composition over a wide range from about 70% C and about 5% H in the lignitic coals to about 93% C and about 4% H in the anthracite coals. The differences are made up of N (< 1%) and S (< 2%) with the remainder occurring as O. In contrast, most petroleums fall into the range 83–87% C



Comparison of the atomic H/C ratios for various fossil fuels.

and 10–14% H with 2–5% S and the remainder occurring as N (< 1%) and O (< 1%). Natural gas which is mostly methane has about 75% C and about 25% H.

Heating value. One important aspect of the fossil fuels is the heating value of the fuel, which is measured as the amount of heat energy produced by the complete combustion of a unit quantity of the fuel. For solid fuels and usually for liquid fuels the heating value is quoted for mass, whereas for gaseous fuels the heating value is quoted for volume. The heating values are commonly expressed as British thermal units per pound (Btu/lb). In SI units the heating values are quoted in megajoules per kilogram (MJ/kg). For gases, the heating values are expressed as Btu per cubic foot (Btu/ft^3) or as megajoules per cubic meter (MJ/m^3). The **table** gives heating values of representative fuels. *See* ENERGY SOURCES; HEAT.

James G. Speight

Bibliography. W. Francis and M. C. Peters, *Fuels and Fuel Technology*, 2d ed., 1980; A. Markuszewski and B. D. Blaustein (eds.), *Fossil Fuels Utilization: Environmental Concerns*, 1986; R. A. Meyers (ed.), *Coal Handbook*, 1981; H. H. Schobert, *Coal: The Energy Source of the Past and Future*, 1987; J. G. Speight, *The Chemistry and Technology of Coal*, 2d ed., 1994; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., 1999.

Fossil humans

All prehistoric skeletal remains of humans that are archeologically earlier than the Neolithic (necessarily an imprecise limit), regardless of degree of mineralization or fossilization of bone, and regardless of whether the remains may be classed as *Homo sapiens sapiens* (anatomically modern humans). In this sense, the term "humans" is used broadly to mean all primates related to living people since the last common ancestor of people and African apes, thus all species currently included in the genera *Homo*, *Australopithecus*, *Ardipithecus*, and *Paranthropus* (and potentially others discussed below).

Discoveries began early in the nineteenth century, although their meaning and antiquity were not recognized before the finding of the Neanderthal specimen in 1856, combined with Darwin's demonstration of the reality of evolution soon afterward. Fossil human remains have come principally from Europe, western Asia, China, Java, and Africa. Because of the rather late entry of humans into the New World, all American remains are of relatively recent origin and recognizable as *H. sapiens*.

Dating fossils. The human lineage was once thought to have come into existence only at the beginning of the Pleistocene, the geological time interval from about 1.8 million years ago to the present. It is now known to have been fully separate from the ancestry of the apes at least as far back as the Early Pliocene or perhaps even the later Miocene (4–7 million years ago, or Ma). This view was widely held by anthropologists, on theoretical grounds, in the early part of the twentieth century. It was abandoned in the 1940s because of lack of supporting evidence. Discoveries in Pliocene and Miocene deposits since then have led to a reappraisal of the evidence (**Fig. 1**).

Dating within the Pleistocene (and earlier) is accomplished by the methods of conventional paleontology and geology, by association with human implements, by several chemical and physical tests for relative age (such as the fluorine test for accumulation of this element in bone), and by geochronometric methods which provide an age in years. The latter usually involve measuring the presence of radioactive substances such as radiocarbon and radiopotassium as well as other radioactivity-based methods such as electron spin resonance, thermoluminescence, and fission-track dating. All of thesemethods are broadly termed radiometric methods, as opposed to approaches such as amino acid racemization or dendrochronology (tree-ring dating). *See* DATING METHODS; RADIOCARBON DATING.

One of the most important methods of correlation between dated geological sequences is paleomagnetism, which does not itself provide a date because it is basically a boolean system with two states (so-called normal and reversed polarity of geomagnetism) of which all occurrences are indistinguishable. But if the pattern of reversals at any site can be matched against the global paleomagnetic time scale and even approximately dated by means of geochronology or paleontology, a sequence of dates may result. *See* PALEOMAGNETISM.

Prehuman ancestry. Humans are catarrhine primates, part of a group including Old World monkeys, apes, and various extinct forms. Most evidence from both comparative morphology and molecular studies of proteins shows that humans' closest living relatives are the African apes: the chimpanzee and the gorilla. Less close is the Asian orangutan, and most distinctive of all apes are the gibbons. A classification which conforms to these relationships within the Hominoidea (apes, humans, and close extinct relatives) recognizes the family Hylobatidae for the gibbons and Hominidae for humans and great apes. The latter family is divided into Ponginae (orangutans and extinct relatives) and Homininae. Many workers, however, continue to reject this view and place humans in the Hominidae as contrasted with the paraphyletic (multi-origined) apes in Pongidae. A few others have suggested that in fact orangutans are the closest living relatives of humans. No fossils of any modern ape are known, other than Pleistocene gibbon and orangutan teeth from southern Asia, and a few Kenyan teeth identified as chimpanzee which were recognized in 2005. *See* FOSSIL APES; FOSSIL PRIMATES; MOLECULAR ANTHROPOLOGY; MONKEY; PRIMATES.

The oldest certain representatives of the Catarrhini are fossils from the Fayum beds of northern Egypt dated around 34 Ma. The best known is *Propliopithecus* (=*Aegyptopithecus*) *zeuxis*, a species near the common ancestry of apes, humans, and Old World monkeys. Hominoids originated in Africa, probably during the Late Oligocene and Early Miocene (26–17 Ma), as suggested by the

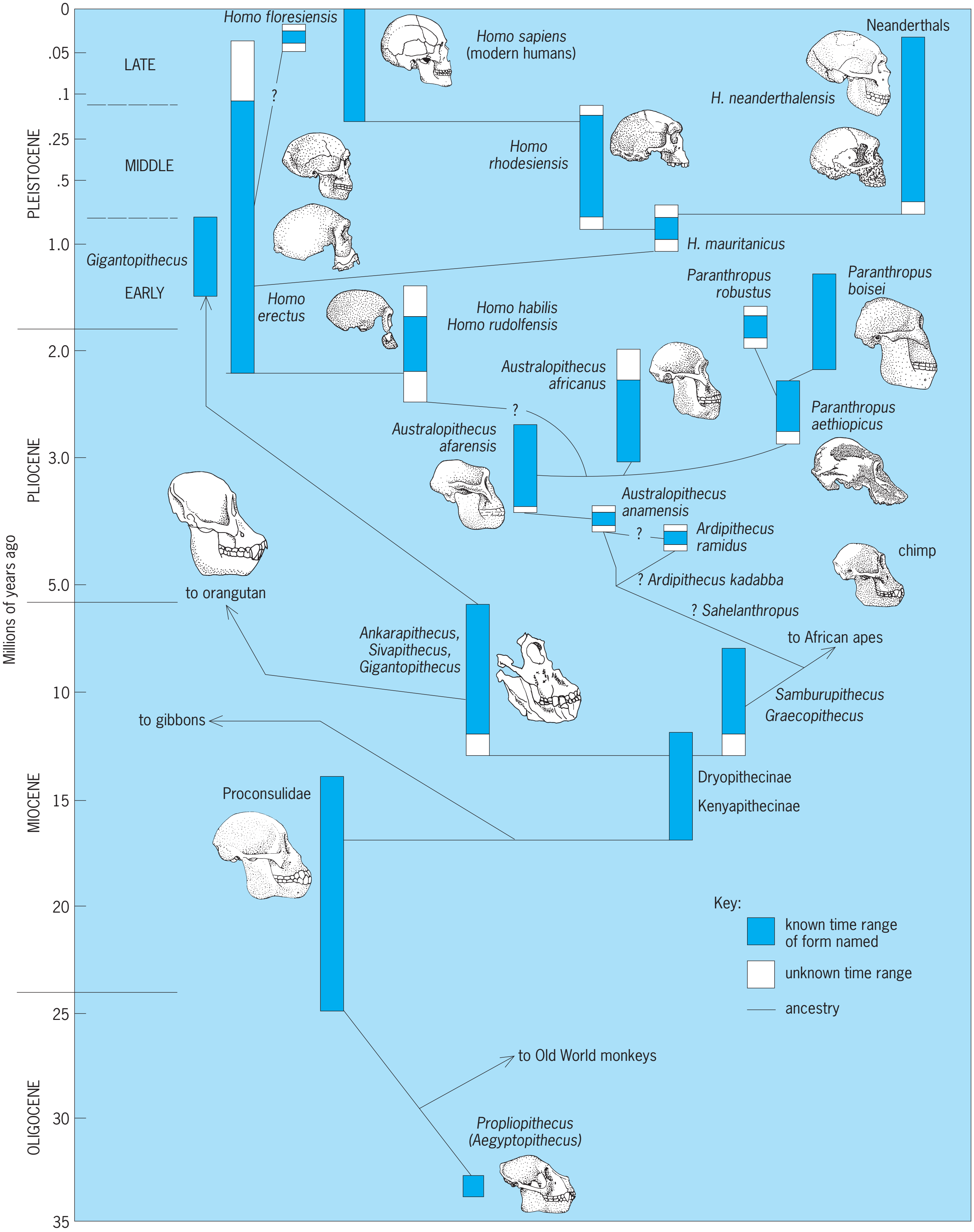


Fig. 1. Human phylogeny from the Oligocene to the present time, showing the skulls of the major known fossil relatives and possible ancestors of modern humans.

genera *Kamoyapithecus* and *Proconsul*, placed in the family Proconsulidae. However, some scholars have questioned the view that these forms are already hominoids, proposing instead that they were conservative catarrhines that predate the evolutionary split between hominoids and Old World Monkeys. Between 20 and 15 Ma, there are several more derived ("advanced") African hominoid genera (for example, *Morotopithecus*, *Afropithecus*, and *Kenyapithecus*) and one Eurasian form (*Griphopithecus*). If these are more closely related to living great apes than are gibbons, as has been proposed in the past, they may represent early members of the Hominidae. However, if they are conservative ("primitive," plesiomorphic) by comparison with the gibbons, then they should not be included in the hominid family.

In the latter case, several workers have argued that the origin of modern ape (and human) lineages must have occurred in Eurasia, where a number of derived genera occur between 14 and 8 Ma. These include *Dryopithecus* and *Pierolapithecus* (discovered in 2004), which were probably close to the common ancestor of all great apes and humans; *Ankarapithecus* and *Sivapithecus*, which appear to represent two early stages in the orangutan lineage (Ponginae); and *Graecopithecus* (**Fig. 2**; also termed *Ouranopithecus*), which has been suggested as a member of Homininae, that is on the lineage leading toward African apes and humans. Another early hominine is *Samburupithecus*, known only from an upper jaw fragment found in the Samburu Hills of Kenya and (like *Graecopithecus*) dated to about 9.5 Ma. A possible alternative interpretation is that a form such as *Morotopithecus*, with relatively derived but poorly known postcranial elements, might have given rise both to the early Eurasian hominids and to an as yet unknown African line which culminated in *Samburupithecus*. In either case, it is not until the latest Miocene (7–5 Ma) that we see the first possible fossil representatives of the Hominini (hominins), the group including modern humans and all our relatives since the last common ancestor shared with chimpanzees. Molecular anthropological studies of DNA sequences and other lines of evidence have suggested that chimpanzees are more closely related to humans than either is to gorillas. This would appear counterintuitive, given the morphological similarity between the two African apes, both of which are also apparently derived in their knuckle-walking locomotion. But similarities may be merely holdovers from the common ancestor of all three forms, while chimps and humans shared an intermediate common ancestor after the gorilla lineage split away. Paleontological evidence is sparse in Africa between 12 and 5 Ma, but it is possible that *Samburupithecus* (or *Graecopithecus*) represents a species near the base of that three-way split or already on the gorilla line. Their ages of 10–9 Ma fit reasonably well with the "molecular clock" estimates of 10–7 Ma for that divergence. The human–chimp split is molecularly estimated at 8–5 Ma or even less, but the human fossil record refutes a date younger than perhaps 5 Ma at the minimum. *See* FOSSIL APES; FOSSIL PRIMATES; MOLECULAR ANTHROPOLOGY; MONKEY; PRIMATES.

Fig. 2. Face (cast) of *Graecopithecus freybergi* from Xirochori, Greece, about 9.5 million years old. (*Courtesy of and © E. Delson; photo by L. Meeker*)

Identification of early hominins. When paleoanthropologists find early fossils that might be on the human lineage, they must identify them as such by locating bony features which distinguish the earliest hominins from their ape relatives or ancestors.

In order to find such features in fossils, we must begin by determining what makes modern humans different from living apes. We can then discern the sequence in which such features appeared in the human fossil record. However, not all changes happened at the same time. This pattern is known as mosaic evolution, referring to mosaic pictures made up of numerous small parts which fit together to make a complete image.

One major difference involves the mode of locomotion used by hominins versus apes. Hominins are habitually bipedal animals (walking on two legs rather than four), whereas our ape ancestor likely included quadrupedal, suspensory and/or knuckle-walking behaviors in its locomotor repertoire. The human skeleton has been modified in many ways to accommodate this change. For example, hominins have a foot with a longer big toe that is in line with the other toes, rather than a grasping ape-like foot with a divergent big toe. The foot has also developed two arches, one longitudinal and one transverse (along the long and short axes of the foot, respectively) that act as shock absorbers when walking bipedally. While the spine of apes is C-shaped, hominins evolved thoracic and lumbar curvatures, resulting in an S-shaped spine that acts to balance the torso over the pelvis and legs. This spine also connects to the skull more vertically (and centrally) underneath rather than toward the back of the skull (as in apes), which can be seen by the more forward placement of the foramen magnum (opening where the spinal cord enters the skull). The pelvis and pelvic muscles have also gone through a major reorganization: the pelvis is less tall and more bowl-shaped, and the gluteus maximus muscle is more developed to help extend the leg backward in walking. The femur (bone of the thigh) is more angled in at the knee in order to center the feet beneath the body, and the knee is capable of straightening out completely. Modern humans also have quite long legs and only moderately long arms compared to body size. All of these features aid in more stable and efficient bipedal locomotion in hominins than apes. The fossil record suggests that relatively modern-looking spinal curvature, pelvic shape, and femoral angling appeared early, while foot arching, perhaps big toe position, and especially fully modern limb proportions did not evolve until later in the hominin lineage. The human skull has also undergone changes not directly related to bipedal locomotion, such as an increase in brain size (relative to body size) and changes in the teeth, including a decrease in canine size. While some early hominins have relatively larger brains than chimpanzees of similar body weight, major brain expansion probably did not occur until the earliest member of our own genus *Homo* evolved.

The next question then is: When in the fossil record do we first see evidence of these changes? Recently discovered fossils from Chad (termed *Sahelanthropus*; **Fig. 3**) may include the oldest hominin cranium with smaller canines and a forwardly positioned foramen magnum (which may in turn suggest some bipedalism). Limb bone fossils from Kenya (named *Orrorin*) are said to have a number of characteristics consistent with bipedalism, though some debate surrounds these claims. Both sets of fossils probably date from 7–6 Ma, making them the oldest possible hominins in existence. Younger finds (∼4.4 Ma) from Ethiopia (*Ardipithecus*) have much smaller canines and may also have a foramen magnum that is positioned more forward than in apes. Canine reduction may have begun even before hominins appeared, as seen in the fossil ape *Graecopithecus*. Further decreases in size eventually led to the very small canines present in modern humans. It has been suggested that smaller canines (with no significant difference between those of males and females) reflect a different social organization, because large canines are associated in apes (and monkeys) with a high degree of competition among males for mates. Fossil australopiths (see below) show clear evidence of bipedalism both in their skeletons and in fossilized footprints found at the site of Laetoli, Tanzania. These hominins still display many apelike features in their skull, showing that habitual bipedalism evolved before modern skull anatomy.

Miocene Hominini. In 2002, French and Chadian researchers led by Michel Brunet reported the discovery of new fossils from Chad in Central Africa. Based on the associated mammalian fossils, the site of Toros-Menalla was estimated to date about 7–6 Ma. Previously, almost all early human remains had been found in eastern or southern Africa, although a few fragments from Chad dating to about 3.5 Ma had been known since the 1990s. The new fossils (named *Sahelanthropus tchadensis*) included a fairly complete though crushed cranium (the skull without the lower jaw) of what might be the oldest hominin ever found. Recent computer-based analysis has permitted a "virtual" reconstruction of what the cranium might have looked like before it was crushed (Fig. 3). Although the brain size is no bigger than that of a similarly sized ape, its teeth are more like a hominins, especially in having an apparently reduced canine. The relatively anterior position of the foramen magnum suggested to the discoverers that this species may have been at least partly bipedal, but others have questioned this view. *Sahelanthropus* is likely near to the chimpanzee–human split, based on both its morphology and the molecular divergence date discussed above.

Fig. 3. Cranium of *Sahelanthropus tchadensis* from deposits ca. 6–7 million years old in Chad, original fossil (left) and virtual reconstruction. (*Courtesy of and © M. P. F. T.*)

In late 1999, Paris-based Martin Pickford and Brigitte Senut working in Kenya discovered a small collection of fragmentary fossils, mostly limb elements, which they also claimed as the earliest known hominin. The fossils were first termed "Millenium Man," as they were announced to the public just before 2000, but they have been formally named *Orrorin tugenensis*. These remains are better dated than those of *Sahelanthropus*, to about 6 Ma, by radiometric techniques. It has been claimed that the femur shows features indicating bipedalism, but most researchers have questioned this interpretation, and some have suggested that *Orrorin* is actually the long-sought fossil of an early African ape ancestor. Unfortunately, dental and cranial remains are too fragmentary to be certain either way.

Still younger fossils are known from sites in the Middle Awash Valley in Ethiopia. Recovered between 1997 and 2004 and first described in 2001, these were named *Ardipithecus kadabba* in 2005. The fossils range in age between 5.8 and 5.2 Ma, at the very end of the Miocene epoch. Most are fragmentary jaws, but a single toe bone has been claimed to indicate bipedalism, although (as usual), some researchers have questioned this interpretation.

Pliocene Hominins: the australopiths. Pliocene humans have been grouped in various ways, but it now seems that four main types can be distinguished. Three of these, dating from roughly 4.5–1 Ma, have previously been assigned to the genus *Australopithecus* and can informally be termed australopiths (here they are placed in the genera *Ardipithecus*, *Australopithecus*, and *Paranthropus*). The fourth group includes early species of *Homo*, beginning about 2.5 Ma. The australopiths have previously been divided into gracile and robust varieties, but the former term really is not accurate and is not used here. All australopith species appear to share a number of basic characteristics (see above) distinguishing them from living and fossil apes and also from later humans, although clearly linking them to the latter. The species of australopith also broadly share a smaller body size than modern humans with surprisingly little difference among the known forms. Estimates of body size based on analysis of weight-bearing joint surfaces and bone lengths suggest average weights around 130 lb (59 kg), with females often under 77 lb (35 kg) and males over 155 lb (70 kg). This quite high sexual dimorphism is also typical of australopiths, as are back teeth that are large for the estimated body size, although the actual sizes and tooth proportions are among the features distinguishing the several species from one another. In light of the diversity now recognized, many researchers accept a division into the genera *Australopithecus* and *Paranthropus*, as well as *Ardipithecus*, but others continue to recognize only one or two of these genera. *See* AUSTRALOPITHECINE.

The fossils of these early humans were first found in South Africa in 1924, but the most recent major discoveries and the best evidence of their age come from East Africa. The South African fossils come from six main sites, which are the remains of ancient cave systems. The australopiths did not live in caves, but their carcasses may have been washed or dropped into crevices near the ground surface by leopards or other carnivores. In two sites especially, many fossils are known from relatively short spans of time (perhaps less than 200 thousand years at each), but the evidence for dating is not definite.

In 1959 Mary Leakey and L. S. B. Leakey discovered a nearly complete australopith skull at Olduvai Gorge, Tanzania, and colleagues dated it at about 1.75 Ma, far older than previously thought. Since then, American and joint American-French and American-Ethiopian expeditions have found hundreds of human fossils in the Omo and Afar areas of Ethiopia, while a Kenyan team has worked around the shores of Lake Turkana (formerly Lake Rudolf), Kenya, and various studies have continued at Olduvai and nearby Laetoli. These regions have yielded smaller numbers of specimens at many separate subsites, but the age of each site can usually be estimated closely by potassium-argon and paleomagnetic dating. Also, many specimens are more complete and show less distortion than their contemporaries from South Africa. Until 1993, no definite examples of *Australopithecus* or *Paranthropus* were known outside these areas, although some claims have been made. The recovery of australopith specimens from Chad (and of *Paranthropus* and early *Homo* from Malawi) demonstrates that new productive regions remain to be explored. It is not clear whether australopiths made stone tools, although they were probably collectors of plant foods, using unpreserved wooden sticks and skin bags, and perhaps scavengers of small game animals.

Ardipithecus: still a question. The oldest known likely human species, *Ardipithecus ramidus*, is known from a small group of fossils found at the Aramis locality in the Middle Awash Valley, Ethiopia. (Additional remains of the species have been reported from the nearby Gona locality as well.) The Aramis fossils have been recovered since 1993 from rocks just above a layer dated to 4.4 Ma, which also yielded animal and plant fossils suggesting a relatively wooded environment. The remains of *Ard. ramidus* include teeth, part of a skull base, and the almost complete long bones of a forelimb, as well as many pieces of a fragmented adult skeleton, including parts of almost all regions of the body. As of early 2006, this skeleton had not yet been formally described or illustrated, but the skull is said to be crushed flat and fragmented into hundreds of tiny pieces which have not yet been reconstructed. The elements described so far document a mosaic pattern combining features similar to those of younger humans (for example, the anterior position of the foramen magnum and nearby structures, implying upright posture and presumably bipedalism; and partial reduction of the canines) with others reflecting retention of apelike conditions (for example, thin enamel covering on molars and incisors; lack of a second cusp, or metaconid, on the lower anterior premolar; relatively large size of canines; and apelike shape of the lower anterior deciduous premolar or "milk molar"). It has been suggested that their presence in a wooded environment implies that australopiths may have differentiated from apelike ancestors in forests rather than open savannahs. But an alternative view suggests that this species represents a "failed" human lineage, one which returned to the forest and secondarily developed thinner enamel convergently with African apes. The remaining bony features of *Ard. ramidus* are all reasonably interpreted as ancestral conditions, to be expected in an ancient human ancestor. Analysis of the partial skeleton should permit determination of the locomotor abilities of this species. Other specimens from Kenya might be referable to this species but are too fragmentary to be sure yet.

Earliest definite humans: Australopithecus. The first habitual bipeds, *Australopithecus*, appear in the fossil record in quantity around 4.2 Ma, during the Early Pliocene. The earliest representatives of *Australopithecus*, named *Aus. anamensis*, have been found since 1994 at sites in the southern Lake Turkana region of Kenya. From Kanapoi, in a layer dated to 4.2–4.1 Ma, come several jaws that differ in shape from those of other *Australopithecus* species. A partial ankle joint from a younger layer confirms upright walking around 4 Ma. It seems likely that *Aus. anamensis* may be close to the ancestry of later species and a possible descendant of *Ard. ramidus*.

Fossils from sites in Ethiopia and Tanzania reveal far more details about a still younger species, *Aus. afarensis* (**Fig. 4**). The most complete material is known from the Ethiopian site of Hadar, about 50 km (30 mi) north of Aramis, where deposits yielded fossils dating between 3.4 and 2.9 Ma. In 1974 a partial skeleton was found and identified as a female by its pelvic bones (and small size compared to other fossils) and nicknamed Lucy. This individual would have stood only 3.5 ft (106 cm) tall and weighed perhaps 65 lb (30 kg). The leg bones of this skeleton indicate that Lucy's legs would have been rather short for a modern human of comparable body weight, but they were in proportion for her estimated stature; this observation supports other evidence that most australopiths were more robustly built than modern humans. Larger males, perhaps 130–175 lb (59–80 kg) in weight, are known from more fragmentary remains, although a nearly complete skull (not known for Lucy) was found in the 1990s and analyzed in detail in 2004 (Fig. 4). Additional specimens, including partial skulls, continue to be found at Hadar. In addition, mandibles and postcranial elements from Maka (across the Awash River from Aramis) and Laetoli (Tanzania) dating around 3.4–3.7 Ma confirm the sexually dimorphic but taxonomically unified nature of the species, while a frontal bone from Belohdelie (3.8 Ma, near Maka) and teeth from the Lake Turkana Basin (3–2.7 Ma) define its temporal range. A fragmentary lower jaw from Chad is estimated to date to a similar time interval by comparing the animal fossils associated with it to those from East Africa; it has been given a new species name but could be a western representative of *Aus. afarensis*.

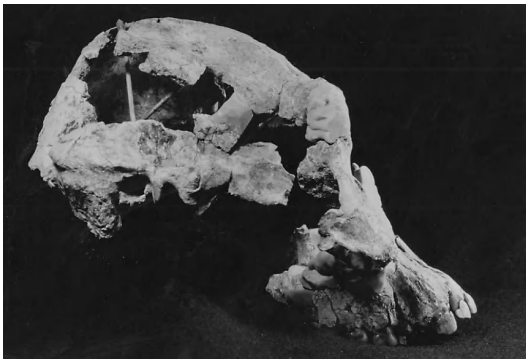


Fig. 4. Cranium of male *Australopithecus afarensis* from Hadar, Ethiopia, dated to about 3.1 million years ago. (*Courtesy of and* © *Institute of Human Origins*)

Lucy's pelvis and leg bones, as well as remarkably preserved footprints from Laetoli, clearly demonstrate that upright bipedal walking was well developed by 3.6 Ma, along with a brain somewhat larger than in modern apes of similar body size. Brain size in *Aus. afarensis* might have been between 350 and 450 ml, in a body weighing some 65–154 lb (30–70 kg) as compared to 365 ml in a 100-lb (45-kg) chimpanzee, 500 ml in a 300-lb (135-kg) gorilla, or 1400 ml in a 150-lb (67-kg) living human. Arguments as to the priority of brain expansion or walking ability in human evolution thus have yet to be fully resolved.

Australopithecus afarensis combines both of these advanced, human characteristics with numerous other features reminiscent of later Miocene hominids and modern apes. The lower face is rather projecting (prognathic), the canines project slightly beyond the level of the neighboring teeth, the anterior lower premolar sometimes has only one major cusp, the tooth row is elongated and nearly parallel-sided, and the forehead is low and retreating. However, the mastoid region (below and behind the bony ear opening in the skull) projects inferiorly more than in either living apes or modern humans. Moreover, the long curved fingers and toes support the idea that it often climbed trees, perhaps to sleep, although there is still some controversy over the degree of posible arboreal behavior in this species.

The geologically youngest of the early australopiths is actually the first to have been recognized. *Australopithecus africanus* was named by R. Dart in 1925 on the basis of a juvenile specimen from Taung (South Africa). It appears to have lived from 3 to 2.3 Ma, but so far it is known only from four sites in South Africa; by far the most numerous specimens come from the middle levels (Member 4) of Sterkfontein cave, dated about 2.8–2.6 Ma (**Fig. 5**). It may have stood 4–5 ft (120–150 cm) tall, weighed 65–150 lb (30–68 kg), and had an average brain size of some 450 ml. The skull seems more lightly built than in *Aus. afarensis*, with a rounded vault but more projecting face. The teeth are more humanlike as well, especially in the presence of two cusps on all anterior lower premolars (like all younger hominins) and less projecting canines. Although the absolute tooth size of *Aus. africanus* is nearly equal to that of small gorillas, the proportions are human, with a smooth decrease in size from molars through

Fig. 5. Cranium of adult (female?) *Australopithecus africanus* from Sterkfontein, South Africa, about 2.5 million years old. (*Courtesy of I. Tattersall*)

incisors. Its postcranial skeleton appears quite similar to that of *Australopithecus afarensis*, with a fully bipedal (if by no means modern) locomotor adaptation. Thus, it was perhaps little changed from its putative ancestor, but those few changes are roughly in the direction of later humans. Some workers have suggested that it might lie close to the ancestry of the genus *Homo*, but pending clearer evidence for its phyletic position, that step is not taken here.

In 1995 researchers described bones of a partial foot from the lower levels (Member 2) of Sterkfontein, perhaps contemporaneous with *Aus. afarensis*. These bones were interpreted as indicating a foot partially adapted to grasping and perhaps tree climbing rather than only to bipedal walking. In late 1998 the find of additional parts of the same skeleton was reported, including both legs, some arm bones, and what appears to be most of the skull. All of these elements must be fully removed from the encasing rock and reconstructed before they can be accurately analyzed, but much more of the skeleton may be preserved, in which case it might surpass Lucy in its completeness. Dates between 4 and 3 Ma have been suggested by different techniques (and some scholars even think it might be younger than 2.5 Ma), and it has not yet been determined if it should be identified as *Aus. africanus*, *Aus. afarensis*, or a new species of australopith.

In 1999, Ethiopian and American researchers described *Aus. garhi* from deposits in the Awash valley of Ethiopia, dated to about 2.5 Ma. A partial face and uncertainly associated limb bones were said to represent a new species close to the origin of *Homo*, but other workers have suggested alternative interpretations, such as that these fossils represented a late population of *Aus. afarensis* or female individuals of *Par. aethiopicus*. Stone tools and animal bones with stone tool cut marks were also found in the vicinity, raising the possibility that australopiths may have manufactured stone tools. On the other hand, early members of our own genus, *Homo*, first appear in the fossil record at other sites around this time period and may have been responsible for making and using these tools.

Robust varieties. Until 1986, robust australopiths were known from two forms, usually accepted as distinct species and increasingly given generic status as *Paranthropus: P. robustus* in South Africa and the more extreme *P. boisei* in East Africa. These species lived between about 2.3 and 1.4 Ma (the age of *P. robustus* is known with less certainty, perhaps 1.9–1.6 Ma) and are distinguished from the other australopiths by their larger size and craniodental specializations. They may have been 4 ft, 6 in. to 5 ft, 9 in. (135–175 cm) tall, weighed 80 to 190 lb (36 to 86 kg), had a somewhat heavy and muscular body build, and a brain size of about 525 ml. The skull is robust, with deep cheekbones and thick lower jaw and often a slightly raised sagittal (midline) crest in the middle part of the skull roof from back to front. These features indicate strong chewing muscles and perhaps a diet of tough foods. The teeth themselves are distinctive: the back teeth (molars and premolars) are large to huge; the front teeth (incisors and canines) are quite small and run nearly straight across the front of the mouth. This difference from other hominins, including apes (which generally have large front teeth and small back teeth), combined with a low forehead and a concave, nearly upright face, further suggests adaptation to powerful chewing. The anterior teeth were probably used as much for grinding as for cutting. The scrappy postcranial elements reveal no major adaptive differences from those of *Australopithecus*.

An even more ancient species, (*P. aethiopicus*), presents a combination of many of these *Paranthropus* features plus others that appear to be holdovers from an *Aus. afarensis*–like ancestry. Here, large molars lie behind sockets for rather large incisors and canines; a concave upper face with low forehead sits above a projecting snout; and the sagittal crest extends back to meet the large nuchal (neck muscle) crest (**Fig. 6**). The brain size was small, near 400 ml. The combination indicates that this species emphasized both large front and back teeth, occupying an evolutionary position intermediate between *Aus. afarensis* and the two previously known robust australopiths; its large mastoid area also links it to *Aus. afarensis*. It is also intermediate in time, ranging from 2.7 to 2.3 Ma.

Australopith relationships. The widely accepted view before 1978 was that *Aus. africanus* represented the common ancestor of the robust forms and *Homo*. The discovery of *Aus. afarensis* led to placement of this species as the basal hominin, with several alternative views of its descendants. In the 1990s the newly recovered *Aus. anamensis* was usually ranked as a still older common ancestor, with *Ard. ramidus* held off to the side, in suspense. Early analyses of *P. aethiopicus* considered it as close to the common ancestor of *P. boisei* and *P. robustus*, but some later studies argued that the three robust species were "only" linked by their common possession of a heavily built chewing apparatus, which might have evolved convergently in two or more disparate lineages. These studies proposed (as had earlier workers on other grounds) that *Aus. africanus* might have been ancestral to *P. robustus* in South Africa or that *P. aethiopicus* might have been an early experiment unrelated to later robust species. Such views seem

Fig. 6. Cranium of adult (male?) *Paranthropus aethiopicus* from West Turkana, Kenya, dated to about 2.5 million years ago. (*Courtesy of A. Walker*)

poorly founded, for if the three robust species are in fact linked by the adaptive complex of "heavy chewing" composed of a number of closely similar elements in each form, it is far more likely that they developed from a single common ancestor than that they were convergently comparable but unrelated. Here the robust clade or lineage is recognized as an evolutionary unit, the genus *Paranthropus*, which was thus long-lived but seems to have had no later descendants.

No known australopith clearly shows features that link it to later human species of the genus *Homo*, but many workers have suggested that *Aus. africanus* might represent the closest approach to such an ancestry yet recovered. However, some have thought that several shared features of the skull point to a common ancestry of *Homo* and *Paranthropus* to the exclusion of *Aus. africanus* (see also below). If *Aus. afarensis* is eventually determined to have been (close to) the common ancestor of both *Paranthropus* and *Homo* (plus *Aus. africanus*?), then it will probably be wise to provide a new genus name for *afarensis* (and also *anamensis*), but for the moment all three species are retained in the genus *Australopithecus*. *See* AUSTRALOPITHECINE.

The rise of *Homo*. The only other genus of the Hominini is *Homo*, true humans, into which all later forms are placed. To summarize in advance, in this essay we suggest that after the brief occurrence of two or three early species of *Homo* in Africa, *H. erectus* evolved about 2 Ma and soon spread into Eurasia east of continental Europe; its far-flung populations were connected by periodic population movements spreading genetic modifications and keeping the species relatively unified. *Homo erectus* persisted, at least in eastern Asia, until 0.2 Ma or later, but farther west around 1 Ma, a geographically restricted species arose near the Mediterranean, with possible representatives in southern Europe and northern Africa. This new species in turn gave rise to at least two major lineages: one in Europe led to the Neanderthals, while a second in Africa continued with little change for half a million years. By about 200 Ka (Ka = thousand years ago) our own species *Homo sapiens* arose in Africa and slowly spread across that continent and into Eurasia, where it eventually replaced Neanderthals and all other human forms.

Early Homo. The identification of the earliest specimens of *Homo* is a subject of debate among paleoanthropologists. In the late 1970s the scientific pendulum had swung back to an idea proposed on less secure grounds by L. S. B. Leakey and colleagues in 1964. They named the species *H. habilis*, based on several finds from Olduvai. Especially significant was the discovery of the remains of a juvenile's lower jaw, with teeth much like those of *Aus. africanus*, and its partial skull, with an estimated cranial capacity of about 685 ml, dated about 1.8 Ma. After much argument over the "reality" and distinctiveness of the new species, it was made clear from additional finds at Olduvai, Lake Turkana, and probably a younger level at Sterkfontein (**Fig. 7**) that a relatively small-brained (510–700 ml) and small-toothed *Homo* was present during 2.0–1.5 Ma. This was thought to be younger than *Australopithecus*, older than (most) *H. erectus*, and contemporaneous with *P. boisei*. A partial skeleton discovered at Olduvai in 1986 has similar teeth and an estimated body size comparable to Lucy; this small size was reported as remarkable but should have been expected, given the similarity in skull size to those of the smaller australopiths.

Several fossils, especially from the Lake Turkana region, appeared to represent a different "morph" or structural pattern. These were typified by skull KNM-ER 1470 (its catalog number in the Kenya National Museum) which has a relatively flat but protruding face, a brain size of about 750 ml, a high rounded vault, and probably large teeth (the crowns are broken off). This (and a few more fragmentary specimens) were at first suggested to represent male individuals, while the smaller individuals were considered females of the same species. Leg bones suggesting a size of 4 ft, 9 in. (155 cm) and 110 lb (50 kg) were originally thought to go with these cranial parts, but it later became clear that early *H. erectus* overlapped in time with the more archaic fossils and had comparable leg bones. Unassociated

Fig. 7. Partial cranium of probable *Homo habilis* from Sterkfontein, South Africa, about 1.9 million years old. (*Courtesy of the late A. R. Hughes*)

postcranial elements have had to be set aside pending the recovery of more bones clearly linked to identifiable crania.

One species or two? These two sets of early human fossils pose an active problem in paleoanthropology: either *H. habilis* had greater sexual dimorphism (especially in brain size) than any known primate, or two (rather homogeneous) species were masquerading under a single name, with the smaller set (the supposed females) most similar in facial morphology to both *Aus. africanus* and later *Homo*. There are at least two opposing solutions to this problem. One view holds that all the known specimens from Olduvai, the Turkana Basin, and South Africa represent the single species *H. habilis*, a larger-brained and bigger creature than *Aus. africanus*, but one with essentially the same dental apparatus—another example of mosaic evolution. It would have had as much sexual dimorphism as did *Aus. afarensis*, perhaps at least as much as in modern gorillas or orangutans, in both form and size of teeth, face, and brain.

Another suggestion is that almost all of the Olduvai fossils, the smaller Turkana region specimens and some from South Africa, are recognized as *H. habilis*, while the 1470 specimen and other larger (non-*Paranthropus*) individuals from Turkana are considered as *H. rudolfensis*. This two-species view is accepted here. In 1992 a partial mandible was found at Uraha in Malawi that is extremely similar to a Turkana *H. rudolfensis* specimen. Not only is Malawi geographically intermediate between the known eastern and southern African site regions, but preliminary age estimates based on associated fossil mammals place the find at about 2.5–2.3 Ma, making it one of the oldest representatives of the genus *Homo*. A temporal bone (the lower middle part of the cranium, where the mandible hinges and the ear is housed) of similar age from central Kenya might also represent this species, and some authorities have even suggested that a few Olduvai specimens belong here as well. One new fossil (the lower part of a face with teeth) described in 2003 from the western part of Olduvai shares certain similarities with both the 1470 fossil and the Olduvai lower jaw first included in *H. habilis*. The describers of this new specimen suggested that the name *H. habilis* would thus best be applied to the larger-sized group (including what had previously been called *H. rudolfensis*), requiring a new name for the remaining smaller fossils; the jury is still out. To further complicate matters, some researchers have suggested that 1470 and its closest relatives might be linked to another Kenyan cranium found in 2001 and dated close to 3.5 Ma. Termed *Kenyanthropus platyops* ("flat-faced Kenya man"), this fossil is somewhat "squashed" and its surface bone broken into many small, slightly separated pieces. It might represent a new lineage of Pliocene hominins; that lineage might include 1470, renamed as *K. rudolfensis*; or the new fossil might just be a badly damaged representative of a known species, such as *A. afarensis*. For the moment, it seems wisest to accept just two named early species of *Homo*, with other options awaiting stronger supporting evidence.

Although australopiths were habitually bipedal, limb bones questionably attributed to *H. rudolfensis* suggest a wider pelvis, perhaps to permit the birth of larger-headed (because larger-brained) infants, which resulted also in the leg bones themselves appearing more modern. It is this combination of apparently significant change in the two major human adaptations of locomotion and intelligence that leads most authors to classify these species as *Homo*, but some have placed both of them in *Australopithecus*. Those researchers take the position that the anatomy of *H. habilis* and *H. rudolfensis* suggests their "adaptive strategies" were more like that of *Australopithecus* than other *Homo* species. The problem with this argument is that most other studies support a stronger evolutionary relationship between *H. habilis*, *H. rudolfensis*, and the other members of the genus *Homo*, including *H. erectus*, Neanderthals and modern humans, than with *Australopithecus*. While the former idea has not taken hold widely in the paleoanthropological community, it does suggest that a fundamental shift in adaptive strategies took place after these early *Homo* species.

It is not clear which species of early *Homo* might have been ancestral in turn to *H. erectus*. Both early forms have been found in the 2.0–1.6 Ma time range in the Turkana Basin and probably Olduvai, where they overlapped with *P. boisei* and *H. erectus*. But neither *H. habilis* nor *H. rudolfensis* as known are morphologically very close to early *H. erectus*. The former species differs not only in small brain and body size but also in limb proportions, although its facial morphology is potentially acceptable in an ancestor of *H. erectus*. The brain of *H. rudolfensis* is large, as might be expected in an ancestor of *H. erectus*, but its teeth are large and distinctive, its face is unique, and its limb bones have not been satisfactorily identified. One wonders if there might not be an as yet undiscovered early species of *Homo* which combined smaller teeth and face with a larger brain. Perhaps the upper jaw from Makaamitalu (found in 1994 high in the Hadar sequence) dated about 2.35 Ma might be a step toward fulfilling this prediction.

Early Homo technology. Evidence of archeological activity has been found with or near some of these early *Homo* specimens, mostly flakes and choppers or "pebble tools" of the Oldowan stone-tool industry (see **table**), and also the remains of small animal prey. It is not possible to tell which types of humans made the tools, but at present the oldest evidence is from Ethiopia, in the Gona region (near Hadar and Aramis) and the Omo valley, about 2.5 Ma. The close time and space concordance of Oldowan tools and early *Homo* suggests that at least one of these species was the actual toolmaker; whichever it was may have been an active hunter of small game, collector of plant and insect food items, and scavenger of larger mammals. *See* PALEOLITHIC; PREHISTORIC TECHNOLOGY.

Homo erectus. While *H. habilis* and *H. rudolfensis* apparently were short-lived and relatively rare African species, their likely successor, *H. erectus*, was common, widespread, and long-surviving. The

Old World archeological variants before 10,000 years ago

Technological mode	Description	Paleolithic subdivisions	Selected regional industries[*]
4	Blade cores, to produce scrapers, points, burins; bone harpoons, art	Late (Upper) Paleolithic	Later Stone Age [sub-Saharan Africa, 40 to < 5 Ka (mode 5, microlithic)] Upper Paleolithic of Europe, including Perigordian Solutrean, Magdalenian Aurignacian (Europe, SW Asia?, > 40 to 27 Ka)
3	Widespread use of prepared cores to obtain variety of flake forms, used as scrapers and projectile points	Middle Paleolithic	Chatelperronian (western Europe, 36–30 Ka) Aterian (northern Africa, 100 to 30 Ka) Mousterian (many variants; western Eurasia and northwestern Africa, 200 to 35 Ka) Middle Stone Age (many variants; sub-Saharan Africa and southern Asia, 250 to 40 Ka)
2	Large bifaces (handaxes and cleavers), also simple cores and retouched flakes	Early (Lower) Paleolithic	In late stages, use of prepared cores and wooden throwing spears Acheulean (Africa and southwestern Asia, 1.5 to <0.2 Ma; eastern Asia, 0.8 to <0.2 Ma; Europe and southern Asia, 0.5? to <0.2 Ma)
1	Simple cores and "casual" flakes	Early (Lower) Paleolithic	Zhoukoudian (China, 1 to 0.25 Ma) Various (Europe, >1 to <0.4 Ma) Oldowan (Africa, 2.5 to <1.5 Ma)

[*] Ka = thousand years ago; Ma = million years ago.

first fossils were found in Java in 1893 and termed *Pithecanthropus erectus*. Most of the later finds in China and across Africa were given distinctive generic and specific names, but all are now usually considered local variants or subspecies of the single species *H. erectus*. Some scholars still recognize additional species, especially *H. ergaster* for the earlier East African populations, but here only one species is accepted. The major anatomical characteristics of this form are the following: a body of nearly modern form and proportions below the neck, topped by a low and slightly elongated skull with cranial capacity averaging 1100 ml (with a range of about 800–1300 ml), smaller teeth in a less projecting face than *Aus. africanus* or *H. rudolfensis*, large solid brow ridges, thick cranial bones, and no chin.

African populations. The earliest specimens are probably from East Africa, dating to as much as 1.9 Ma. These populations might have continued to make simple Oldowan artifacts, as no other form of stone tools is known at this time. There is evidence of the use of fire as early as 1.8 Ma and in scattered sites thereafter. Acheulean bifaces (handaxes and cleavers) appear by 1.5 Ma, though only rarely in direct association with *H. erectus* fossils; evidence for group hunting of large animals is rarer still and has been questioned by some scholars. In addition, the eventual extinction of *P. boisei* in this time interval may have been caused by direct or indirect competition with the more advanced *H. erectus*.

The most important African *H. erectus* find of recent decades is a mostly complete skeleton of an adolescent male (its sex judged by pelvic bones as well as by relative cranial robusticity), recovered on the west side of Lake Turkana in 1984 (**Fig. 8**). Judging by the pattern of dental eruption, he was probably 9–10 years old at death, although an age of 11–12 has also been suggested. One still ambiguously answered question is whether this species underwent the adolescent growth spurt which occurs in modern children of ages 12–14. Using formulas based on limb bone lengths of modern human adults and children, it was estimated that the "Turkana boy" might have reached an adult height of nearly 6 ft (183 cm) and a body weight of 150 lb (68 kg). Moreover, his body shape would have been slender or elongate, much as seen in living people of the same region, as a result of similar adaptation to a warm and dry climate. The development of brow ridges and other cranial structures in the adolescent male and a larger skull from Olduvai contrasts markedly with other Turkana skulls, indicating the persistence of strong sexual dimorphism.

A number of interesting new fossils thought to be *H. erectus* were found in East Africa in 1995–2002. A relatively complete cranium without a face from Lake Turkana and cranial fragments from another Kenyan site represent quite small individuals. The latter was found in the same stratigraphic layer as many bifacial Acheulean tools, good evidence that they were indeed the toolmakers. The first potential *H. erectus* fossils from Ethiopia and Eritrea have also been recently described, both from around 1 Ma. Measurements of the Ethiopian fossil indicate its similarity to both African and Asian *H. erectus* fossils. The Eritrean fossil, on the other hand, shows an increase in width of the back of the skull that is also seen in modern humans. Together, these new fossils extend the known variation in African *H. erectus* considerably.

Eurasian representatives. *Homo erectus* (perhaps as a result of increasing population size) spread from Africa into Eurasia through the Middle East, perhaps earlier than has previously been thought. New dates for some long-known Javanese fossils suggested ages around 1.7 Ma, but these have been questioned.

Fig. 8. Partial skeleton of adolescent male *Homo erectus* from West Turkana, Kenya, about 1.5 million years old. (*Courtesy of and © National Museums of Kenya*)

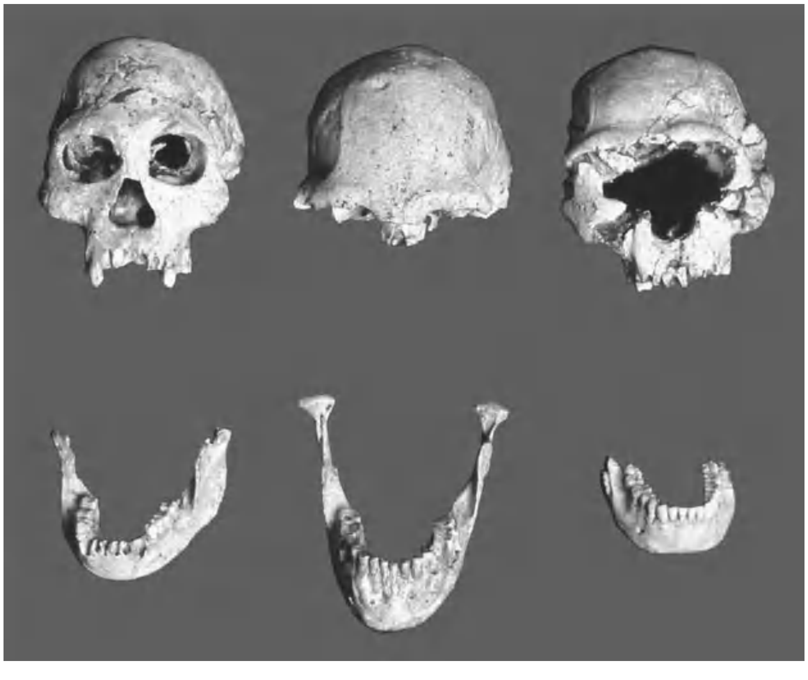


Fig. 9. Crania and mandibles (in front view) of *Homo erectus* from Dmanisi, Georgia, reveal diversity in a single sample dated ca 1.7 Ma. (*Courtesy of and © Georgian National Museum*)

The most exciting new finds have come from the site of Dmanisi (Georgia), mainly from 1999 to 2005 (**Fig. 9**). There are currently five crania, four mandibles and a large number of postcranial elements along with Mode 1 (Oldowan-like) tools and animal bones dated to 1.8–1.6 Ma. Mode 1 and early Mode 2 artifacts from Israel (as yet not definitively found with human fossils) are slightly younger. Thus, *H. erectus* in the broad sense accepted here must have been the first human species to leave Africa in large numbers, probably about 1.8 Ma, soon after it first appeared in the fossil record.

Fossils of this species may extend in mainland Asia to nearly 200,000 years ago, mostly associated with fauna from the warmer intervals in this time of alternating glacial climate. From central Java, Indonesia, come a series of cranial and dental remains spanning from perhaps 1.7 Ma to possibly 30 Ka. The earlier range of time yielded several groups of skulls known popularly as Java man, including the first found members of this species; these specimens (and also those from China) tend to be more derived or extreme in their morphology than most African specimens, including the common presence of a thickened "keel" of bone down the midline of the braincase and a more projecting browridge. A dozen younger partial skulls from Ngandong, Java were long thought to be quite recent in age and once were placed in *H. sapiens*, but later study indicated close similarity to earlier Indonesian *H. erectus*, as they are now classified. Their age remains uncertain, with some evidence suggesting a range between 400 and 100 Ka, while a dating attempt in 1996 arrived at the surprisingly young age of 50–30 Ka, which has been questioned on several grounds. If this is indeed a correct date for these fossils, this would be the first demonstration of the coexistence of *H. erectus* with anatomically modern humans, known elsewhere in Indonesia and Australia at that date. *Homo erectus* specimens continue to be recovered from Java, and one partial cranium was even recognized in 1999 at a New York City shop selling natural history specimens (**Fig. 10**). It was determined that it had been found in Indonesia some years earlier and illegally removed from that country; the shop's owner returned it to Indonesian authorities.

A fascinating find was made in 2003 on the Indonesian island of Flores in a cave called Liang Bua, which has an archaeological record preserved from 200 Ka to the present. The fossils found there, including a well-preserved skull, additional lower jaws, and postcranial bones, have been dated to about 18 Ka. Dubbed "Hobbits" in the popular press, the remains are those of a small-brained (380 ml) and small-bodied (3 ft 3 in. or 1 m) hominin with a skull that resembles that of a small *H. erectus*. This led the discovers to conclude that a population of *H. erectus* somehow reached the island over water and became "dwarfed" after isolation, a phenomenon observed in other large mammals on islands; they named the fossils *H. floresiensis*. Other scientists have suggested that the remains belonged to pathological modern humans, which

Fig. 10. Partial cranium of ?female *Homo erectus* from Sambungmacan, Java, Indonesia, age uncertain (probably between 500,000–100,000 years ago). This specimen was identified in a natural history store in New York City, years after it was found by local workers in Indonesia. (*Courtesy of and © E. Delson; photo by C. Tarka*)

seems unlikely given the number of individuals of differing ages recovered. Further finds published in late 2005 highlighted the long arms and other anomalous postcranial features of this species, which are difficult to explain.

The first Chinese *H. erectus*, once called Peking man, was found at Zhoukoudian, near Beijing, where they occupied a large cave during most of the period between 500 and 250 Ka. Additional fossils of this form are known from Yuanmou in the southwest, Lantian and Yunxian in the center, and Hexian and Tongshan (Nanjing) in eastern China; some may be as old as 1–1.2 Ma. Archeological remains found with these fossils are mainly of the simple Mode 1 variety (no stone tools are unambiguously known from layers with human fossils in Java). However, more complex bifacial tools have been found in the Bose Basin of southern China, from about 0.8 Ma. These tools are comparable in overall form and complexity to Acheulian (Mode 2) tools found in Africa at this time and predate Mode 2 tools in Europe. This contradicts the long held hypothesis that only simple Mode 1 tools were made in East Asia while more complex Acheulean tools were restricted to Africa and Europe at this time. Although there have been claims, no definite *H. erectus* fossils are yet known from Europe, nor are archeological remains or more modern humans unambiguously documented there older than about 800,000 years. The Dmanisi specimens from Georgia are the most ancient human representatives in the region, but the lack of further remains over the succeeding million years may indicate that *Homo* required further physical or cultural adaptations in order to survive in Europe proper.

Middle Pleistocene Homo. It has been suggested that the increased rigor of the glacial climate in Europe at this time was the impetus leading to the evolution of humans who seem to be physically more modern in several ways than Afro-Asian *H. erectus*. These people were often termed early or archaic *H. sapiens*, or sometimes placed in their own species, *H. heidelbergensis*. This view was acceptable so long as the most ancient African representatives of this group were poorly dated or younger, but studies in the later 1990s suggested some modification. The earliest human fossils in Europe were long thought to date to about 500 Ka in England (a tibia from Boxgrove) and Germany (the mandible from Mauer near Heidelberg found in 1908). The long-known human fossils from Tighenif (previously called Ternifine, in Algeria), dated to 800–700 Ka, were transferred out of *H. erectus* by some workers because they present at least one derived feature of the lower jaw. Moreover, the partial skull from Bodo (Ethiopia) was also dated older than 600 Ka, and similar South African fossils were estimated to be of comparable age. These dates suggested that the earliest representatives of these more derived humans may have lived in the northern half of Africa and perhaps evolved there from local *H. erectus* populations early in the Middle Pleistocene (780–125 Ka).

Then, in 1995, fragmentary human fossils and associated stone tools from the older levels at Atapuerca (Spain) were dated to about 800 Ka, implying that archaic *H. sapiens* appeared at about this date all over the western Old World. These specimens were named *H. antecessor* in 1997, and it was suggested that they represented the common ancestor of all later human varieties. Meanwhile, in 1996, a cranium lacking the face was described from Ceprano, Italy, in a context suggestive of a date of about 700 Ka or more. However, this date is not particularly secure. A revised reconstruction claims strong similarity to *H. erectus*, otherwise unknown in Europe. In late 2005, stone tools (of both Mode 1 and Mode 2 types) and animal fossils were described from Pakefield in eastern Britain, a site estimated to date to 700 Ka; no human fossils are yet known, but they are eagerly awaited.

All of these finds, combined with theoretical arguments about the best way to recognize and delimit species in the fossil record, have led to competing interpretations of the number of species of *Homo* known in the past million years. Some workers continue to place all post-*erectus* fossils in “archaic *Homo sapiens*,” sometimes recognizing a variety of temporal and geographic subspecies (such as the Neanderthals and anatomically modern humans). A few have gone so far as to include *H. erectus* within an over-enlarged *H. sapiens*. At the other extreme, some researchers accept between five and seven species in the same time period: *H. antecessor*, *H. cepranensis* (for Ceprano) *H. heidelbergensis* (either restricted to Europe or extended to Africa and even East Asia), *H. rhodesiensis* (for African Middle Pleistocene populations), *H. neanderthalensis*, *H. sapiens* (restricted to anatomically modern humans), and perhaps others. A possible middle ground would be to (1) include the earliest of these African and European populations in one named group [for example, *antecessor*, or perhaps *mauritanicus* (the name originally given to the Tighenif fossils), as suggested by J.-J. Hublin]; (2) combine all post-500 Ka nonmodern European fossils in *neanderthalensis* (including *heidelbergensis* here); (3) group nonmodern African fossils younger than Tighenif in *rhodesiensis*; and (4) restrict *sapiens* to anatomically modern humans worldwide. These "groups" could be given specific or subspecific status, depending on the theoretical model used. Recent work discussed below supports full species status for Neanderthals based on several lines of evidence, and all four groups are here interpreted as full species. While both fossil and genetic evidence supports an African origin for modern humans, the evolutionary relationships among all of these groups remains unclear.

Homo rhodensis and early representatives of *H. neanderthalensis* occur in Africa and Europe, respectively, between 650 and 250 Ka, thus contemporaneous with *H. erectus* populations in eastern Asia. They share somewhat larger brains (for body size), smaller teeth, more expanded facial sinuses and occiput (rear of the skull), but less robustness than in *H. erectus*. All these features are found in more extreme form in modern humans and in the late "classic" Neanderthals (see below). In most areas, these people still used Acheulean tools, but perhaps with greater efficiency. It is likely that these "intermediate-grade" humans spread gradually eastward across the Old World, replacing late-surviving populations of *H. erectus* everywhere by 200 Ka. These geographic variants were not only distinct from *H. erectus* but also from each other to a greater degree than is true among living varieties or "races" of anatomically modern humans. In southern Africa, one cranium was found at Broken Hill, now Kabwe, Zambia (formerly Northern Rhodesia, hence the name Rhodesian man), and broadly similar specimens are known in South Africa (Saldanha and Florisbad), Tanzania (Ndutu), Ethiopia (Bodo), and Morocco (Salé and Thomas quarries). These people made Acheulean or equivalent Mode 2 tools and apparently hunted big game across most of Africa. In early 2006, a new partial cranium probably belonging to this species from Gawis (Ethiopia, near the Gona Pliocene tool site) was reported by Ethiopian and U.S. researchers led by Sileshi Semaw. Its age is as yet uncertain but probably in the 500–250 Ka interval.

Fig. 11. Crania in frontal view of male *Homo neanderthalensis* from La Ferrassie, France (on right) and Amud, Israel, ca 50,000–60,000 years old. (*Courtesy of and © Israel Antiquities Authority (Jerusalem), Musée de l'Homme (Paris) and E. Delson; photo by C. Tarka*)

Rare specimens from China appear to be younger, mainly dating to about 250–150 Ka. These include a nearly complete cranium from Dali, in central China, and a partial skull and skeleton from Jinniushan (or Yingkou), in the northeast, as well as scattered, less complete remains. All of these fossils, especially Dali, are broadly similar to the African specimens just mentioned, as well as to some of the earliest European *H. neanderthalensis*. Other Chinese specimens, such as Maba (from the southeast), and the central Indian Hathnora (or Narmada) fossil, are partial crania which are both younger (perhaps about 150–75 Ka) and more derived morphologically, although not in the direction of either early anatomically modern people or the contemporaneous Neanderthals.

Neanderthals. The best known of the archaic varieties are the Neanderthals, from Europe and western Asia. It now seems likely that this group evolved locally in Europe from early Middle Pleistocene *Homo* via intermediate populations previously called *Homo heidelbergensis* such as those known from England (Swanscombe, ?Boxgrove), Spain (Atapuerca), France (Arago, Montmaurin), Germany (Mauer, Steinheim), and Greece (Petralona). They became adapted to the cold climates of glaciated Europe, with prototypical Neanderthal anatomy well established by about 200 Ka. During the warm interval about 120 Ka, they may have spread into the Near East and central Asia. In the cold glacial phase between 110 and 35 Ka, "classic" (or extreme), cold-adapted Neanderthals were abundant in cold northern parts of western and central Europe, while less extreme forms (perhaps more like their immediate predecessors) inhabited areas to the south and east. They were essentially stocky humans, but had long, low skulls with a projecting occipital region, large faces, teeth, and brow ridges; and brains averaging 1500 ml in volume (**Fig. 11**). Their limbs and trunks were heavily muscled, indicating great strength, but many bones were broken and healed during life. They made Mousterian tools (a variant of Middle Paleolithic or Mode 3 flake-based tool kits), often lived in caves or wooden shelters where they controlled fire, hunted big game, and had primitive religious beliefs, including burial of the dead with grave goods. There is intense argument among paleoanthropologists as to how modern the Neanderthals were behaviorally, in terms of their stoneworking and hunting techniques and modes of foraging, whether planned or merely ad hoc. For example, the frequent broken bones may have resulted from hunting large game at close proximity, rather than using projectile technology from a distance. Such controversies feed back into the question of whether the Neanderthals are a distinct species or a distinctive subspecies of *H. sapiens*. A related question is whether the Neanderthals were in any way ancestral to anatomically modern humans, especially of Europe.

Recognition of a separate Neanderthal species implies an almost absolute reproductive isolation and lack of genealogical continuity, while the opposite is true for most interpretations of Neanderthals as members of *H. sapiens*. The interpretation accepted here is that *H.neanderthalensis* is considered to have been geographically and culturally isolated from early anatomically modern humans and their ancestors, two independent but closely related lineages (species) evolving in parallel until they finally met, after which the former group soon became extinct. Indeed, current research suggests that differences in the skulls between Neanderthals and modern humans appeared early in individual growth, meaning that Neanderthal children looked quite different from modern human children. Neanderthals also differ in developing more rapidly than modern humans. In other words, they achieved their adult anatomy faster than modern humans do. In addition, work by K. Harvati and colleagues published in 2004 shows that the differences in skull anatomy between Neanderthals fossils and modern humans are at the level of separate species when compared to other primates. Furthermore, when scientists extracted mitochondrial DNA from the bones of several Neanderthal fossils they found that Neanderthals have some unique sequences of DNA when compared to modern humans. Neanderthals, possibly due to local climatic differences, were probably isolated from other hominins and as a result became quite distinct in their development as well as their skull shape compared to modern humans. *See* NEANDERTHALS.

Spread of modern humans. One of the major foci of recent paleoanthropological research is the clarification of the area of origin and the early history of modern humans, *H. sapiens*. The skull of this form is characterized by a small, nonprotruding face; small teeth and brow ridges; a chin; and a high, rounded braincase. There are no specimens of this type known (or even hinted at) anywhere in the world earlier than about 200–150 Ka. New fossils representing early modern humans were recovered from Herto, Ethiopia, in 1997 (published 2003) and dated 164–150 Ka. Aside from having larger faces than recent humans, these are quite modern looking. Moreover, in 2005, the long-known fossils from Kibish (Omo Valley, Ethiopia) were dated close to 200 Ka. This confirms previous suggestions that about 200–100 Ka some fossils from eastern and southern Africa (for example, one from Kibish and one from a late horizon at Laetoli) suggest the persistence of a "Rhodesian-like" morphology, while others (for example, another from Kibish and several from Djebel Irhoud, Morocco) appear to be nearly modern. Two somewhat younger sites in South Africa have produced important additional modern human fossils. At Border Cave, a partial cranium and other fragments may date to nearly 90 Ka; they are clearly modern in form, but their date is questionable. The Klasies River Mouth caves, on the southern coast, have yielded a sequence of layers with good dates and archeological context; the human remains dated about 100 Ka are scrappy but appear modern, with a chin, small brow ridges, and overall gracility. Such gracilization, especially of the face, is a major feature of the evolution of (and within) *Homo sapiens*. In combination, these remains and other, less complete fossils indicate that early moderns were living in sub-Saharan Africa by 150 Ka. Archeological remains dated 100–75 Ka in South Africa and D. R. Congo (ex-Zaire) indicate that at least some of these people were making Mode 3 (MSA) tool kits with elements (such as bone harpoons, engraved bone, and minerals and shell beadwork) that do not appear in Europe until after 30 Ka. *See* EARLY MODERN HUMANS; PALEOLITHIC; PREHISTORIC TECHNOLOGY.

From such a possible sub-Saharan origin, anatomically modern *H. sapiens* appear to have spread across the Old World, differentiating into local races by 80–50 Ka. This view of human dispersal has received support from studies of the distribution pattern of human mitochondrial DNA haplotypes (variants) and other genetic evidence. The majority of these studies suggest that the major dichotomy in modern human population genetics is between Eurasians and Africans. Such results fit well with the fossil evidence for African versus Eurasian divergence about 100 Ka. Moreover, dates on early anatomically modern remains from Israel (Djebel Qafzeh and Skhul) documented the presence of the ancestors of Eurasians outside Africa by about 110–90 Ka. This is especially intriguing because most Israeli Neanderthals have been dated to about 65–45 Ka, which is significantly younger than the early moderns and implies that they may have overlapped in this region of the world. Even more complexity is implied by the strong similarity of Mousterian tool kits associated with such diverse human forms as the Neanderthals of Europe and Israel (and farther east), the early moderns from Israel, and the "premoderns" from Jebel Irhoud (Morocco), but the implications of this cultural similarity are as yet unclear.

Found in the late 1970s, the youngest known Neanderthal skull comes from southern France and is associated with tools of the Châtelperronian industry, a Mode 3 or 4 variety previously thought to have been made by *H. sapiens*. This specimen dated to 34 Ka has been alternatively interpreted as the maker of these tools (possibly after contact with Late Paleolithic moderns) or as evidence for direct Neanderthal ancestry of moderns. In some cases, more fossils do not answer questions but create new ones. Sites in southern Spain and Portugal have yielded less complete Neanderthal fossils and Mousterian tools dated about 30 Ka, after which modern *H. sapiens* was the sole form of human to be found anywhere. One reason for the success of *H. sapiens* may have been their greater tool-making efficiency, as documented by the Late (or Upper) Paleolithic Mode 4 blade-and-burin industries. These people included large quantities of worked bone in their tool kits (using chisel-like burins to carve and engrave the bone), constructed dwellings of wood or of already fossilized animal bone, hunted large game, fished with harpoons, and in general behaved much like their living descendants. In many parts of the world, they also wore personal ornamentation (jewelry) and engaged in artistic pursuits, including carving small animal statues and perhaps calendars, as well as painting on the walls of rock overhangs and deep caves. *See* PALEOLITHIC; PREHISTORIC TECHNOLOGY.

In 1999, Portuguese and American paleoanthropologists described the remains of a 4-year-old child, buried near Lagar Velho, Portugal, about 25 Ka. The burial pattern and most of the child's morphology indicated links to the Gravettian culture, made by early anatomically modern people elsewhere in Europe at this time. However, features of the lower leg bone (tibia) and lower jaw (the cranium was crushed and later reconstructed) suggested similarity to Neanderthals. The describers hypothesized originally and in a major publication in 2005 that this individual might have been the result of hybridization between Neanderthals and moderns, but other researchers argued that the morphology was not that different from what could be expected in a robust anatomically modern child and that hybridization would have resulted in features intermediate between Neanderthals and moderns, not clear features of each.

Many names have been given to early modern humans, especially in Europe, but these indicate only minor differences. The term Cro-Magnon derives from several skeletons found in 1868 in Les Eyzies, France. They gave their name to a "race" said to occur either just in France or across most of Europe. In fact, Cro-Magnon people were already essentially Europeans, while early Africans are known from sites in eastern and southern Africa. Australasia was colonized over water after about 70 Ka, with important finds at Keilor and Lake Mungo. New World Indians likely originated from Siberia, by means of crossing a land bridge over what is now the Bering Strait. Many human fossil remains are known in the Americas as far back as 12 Ka, but some dates as old as 35 Ka have been obtained on archeological sites, indicating that perhaps several crossings of the land bridge occurred.

In contrast to the above "Out of Africa" view of human dispersal (based on the idea that modern humans evolved in sub-Saharan Africa more than 100,000 years ago from populations of archaic *Homo*) accepted here, a minority view (the "Multiregional" hypothesis) interprets the fossil record to document the nearly parallel origin of modern humans in different regions of the Old World from a *H. erectus* ancestry. Each regional variety is said to present morphological characteristics linking archaic to modern populations, while gene flow (migration and interbreeding) between regions kept the geographical varieties united in a single species at any one time. While most researchers reject a significant contribution from Neanderthals or *H. erectus* to the modern human gene pool, there may still have been some interbreeding among the species. *See* EARLY MODERN HUMANS; PALEOLITHIC; PREHISTORIC TECHNOLOGY.

Eric Delson; Karen L. Baab

Bibliography. E. Delson (ed.), *Ancestors: The Hard Evidence*, 1985; E. Delson et al. (eds.), *Encyclopedia of Human Evolution and Prehistory*, 2d ed., 2000; J. G. Fleagle, *Primate Evolution and Adaptation*, 2d ed., 1998; W. C. Hartwig, *The Primate Fossil Record*, 2002; R. G. Klein, *The Human Career: Human Biological and Cultural Origins*, 2d ed., 1999; J. H. Schwartz and I. Tattersall, *The Human Fossil Record*, vols. 1–4, 2001–2005; A. Walker and R. Leakey (eds.), *The Nariokotome Homo erectus Skeleton*, 1993; P. F. Whitehead, W. K. Sacco and S. B. Hochgraf, *A Photographic Atlas for Physical Anthropology*, 2005.

Fossil primates

Extinct members of the order of mammals to which humans belong. All current classifications divide the living primates into two major groups (suborders), but zoologists differ as to whether the tarsier (*Tarsius*) should be classified with the lower primates (lemurs, lorises, and bushbabies) or the higher primates (New and Old World monkeys, greater and lesser apes, and humans).

All primates have a common origin that, however, is not reflected in the universal possession of a suite of diagnostic features. The order as a whole has been characterized in terms of showing a group of progressive evolutionary trends, notably toward predominance of the visual sense, reduction of the sense of smell and associated structures, improved grasping and manipulative capacities, and enlargement of the higher centers of the brain. Among the extant primates, the lower primates more closely resemble forms that evolved relatively early in the history of the order, while the higher primates represent a group more recently evolved (**Fig. 1**).

A classification of the primates, as accepted here, is shown on the next page.

Early primates. The earliest primates are placed in their own semiorder, Plesiadapiformes (as contrasted with the semiorder Euprimates for all living forms), because they have no direct evolutionary links with, and bear few adaptive resemblances to, any group of living primates. However, the chewing teeth and the locomotor anatomy of these fossil forms sufficiently resemble those of later primates to demonstrate the common origin of the two groups. Best known from the Paleocene Epoch, around 65–55 million years ago (Ma), and found in both the Old World and the New World, most plesiadapiforms retained clawed hands and feet. However, in 2002 it was shown that at least one species, *Carpolestes simpsoni*, possessed an opposable great toe with a nail, as in euprimates (**Fig. 2**). Plesiadapiforms also had rather small brains compared to their body size, possessed large specialized front teeth, were probably arboreal in habit, and may have been adapted for terminal branch feeding, as recently suggested. These animals are known from fossil deposits as far north as Ellesmere Island, in Arctic Canada, which, during the Paleocene, was covered by the subtropical forest stretching continuously from western North America across a landlocked North Atlantic into western Europe.

Eocene primates. There is no known plesiadapiform that is a satisfactory candidate for the ancestry of the fossil primates of modern aspect typical of the succeeding epoch, the Eocene (55–34 Ma). Often

- Primates
  - Semiorder Plesiadapiformes (archaic extinct primates)
    - Superfamily: Paromomyoidea (paromomyids, picrodontids)
    - Plesiadapoidea (plesiadapids, carpolestids)
  - Semiorder Euprimates (modern primates)
    - Suborder Strepsirhini (toothcombed prosimians and extinct allies)
      - Infraorder Adapiformes (extinct early strepsirhines)
        - Superfamily Adapoidea (extinct early strepsirhines)
      - Infraorder Lemuriformes (modern strepsirhines)
        - Superfamily: Lemuroidea (typical lemurs)
        - Indrioidea (indris, aye-ayes, and subfossil relatives)
        - Lorisoidea (lorises, bush-babies, mouse and dwarf lemurs)
  - Suborder Haplorhini (tarsiers and higher primates)
    - Hyporder Tarsiiformes (tarsiers and extinct relatives)
      - Superfamily: Tarsioidea (tarsiers and close relatives)
      - Omomyoidea (extinct early haplorhines)
    - Hyporder Anthropoidea (higher primates)
      - Infraorder Paracatarrhini (archaic anthropoids)
        - Family: Oligopithecidae (archaic protoanthropoids)
        - Parapithecidae (extinct Oligocene monkeys)
      - Infraorder Platyrrhini (New World anthropoids)
        - Superfamily Ateloidea (New World monkeys)
          - Family: Atelidae (howler, spider, saki, titi, and owl monkeys)
          - Cebidae (squirrel, capuchin, and marmoset monkeys)
      - Infraorder Catarrhini (Old World anthropoids)
        - Parvorder Eocatarrhini (archaic catarrhines)
          - Family: Propliopithecidae (extinct common ancestors of hominoids and cercopithecoids)
          - Pliopithecidae (extinct early catarrhines)
        - Parvorder Eucatarrhini (advanced catarrhines)
          - Superfamily Hominoidea (apes and humans)
            - Family: Proconsulidae (extinct early apes)
            - Hylobatidae (gibbons, lesser apes)
            - Hominidae (great apes, humans, and extinct relatives)
          - Superfamily Cercopithecoidea (Old World monkeys)
            - Family Cercopithecidae (Old World monkeys)

termed euprimates, they are divided broadly into lemurlike forms, usually grouped into the superfamily Adapoidea, and tarsierlike forms (Omomyoidea), although this division may prove to be oversimplified. Eocene primates of both the Old and New Worlds already display the trends that mark modern primates as a whole: These arboreal animals possessed grasping hands and feet in which sharp claws were replaced by flat nails backing sensitive pads; the snout was reduced, suggesting a deemphasis of smell, while the bone-ringed eyes faced more forward, producing stereoscopic vision and suggesting primary reliance on the sense of sight; and the brain was somewhat enlarged relative to body size when compared to those of other mammals of the period. A new fossil skull from China reported in 2004 (**Fig. 3**) documented the presence of omomyoids in the earliest Eocene of eastern Asia, and its relatively small orbits suggested that it was diurnal.

It is possible that the origins of some specific groups of extant lower primates may be traced back to or through certain known primate genera of the Eocene. In 2003, two distinct primate dentitions bearing the toothcomb characteristic of all modern strepsirhines were documented in Egypt to date back to the Eocene epoch, around 40 Ma. It has been suggested that one of these forms is related to the living bushbabies and the other to the lorises. Similarly, some dental fossils reported in 2001 from 30 Ma deposits in Pakistan have been identified as lemuroids. Even if this is not the case, however, it is universally accepted that the antecedents of the living lower primates are to be sought somewhere within the Eocene primate radiation, though the details of this ancestry still remain unclear. In North America and Europe, lower primates had nearly disappeared by the close of the Eocene, while virtually all known fossil lower primates of later epochs in Africa and Asia are closely related to modern primates of these areas.

Modern lower primates. The extant lower primates are allocated to the suborder Prosimii if *Tarsius* is included, and to the suborder Strepsirhini if this strange primate is excluded, as is provisionally done here. All extant strepsirhines possess dental scrapers (toothcombs), in which the lower front teeth are elongated, closely approximated, and forwardly projecting. These specialized structures are used both

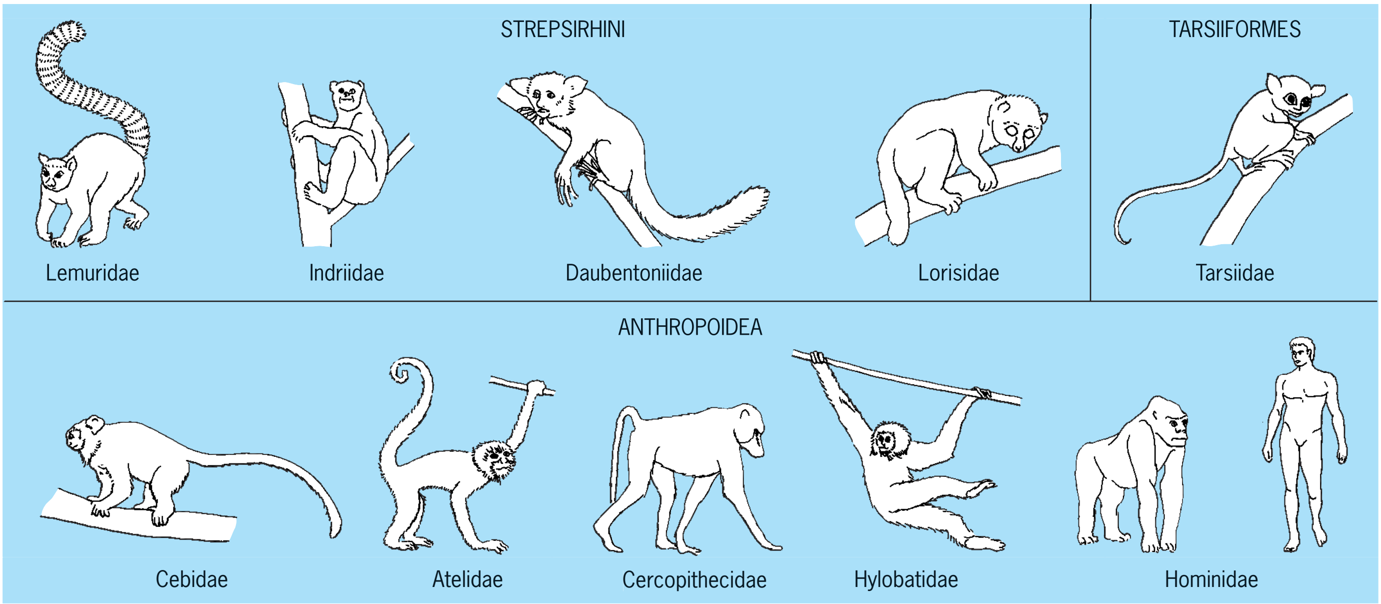


Fig. 1. Representatives of living primate families.

in feeding and in social grooming. Additionally, all living strepsirhines retain a moist, naked rhinarium (wet nose, as in a dog) and related structures, reflecting a greater reliance on the sense of smell than is typical of higher primates. All are also united by possessing a toilet claw (used for self-cleaning) on the second digit of the foot. Although all strepsirhines possess grasping hands and feet, their manual dexterity is generally inferior to that of higher primates; their brains are also relatively small.

There is no ancient (older than 10 Ka) primate fossil record in Madagascar, home of the most diverse group of modern lower primates, but fossils of now-extinct species document a much wider adaptive radiation before the arrival of humans on the island less than 2000 years ago. Notable among those extinct lemurs are *Palaeopropithecus*, a large-bodied under-branch hanger, somewhat slothlike in its locomotion (**Fig. 4**); the semiterrestrial and dentally specialized members of the subfamily Archaeolemurinae, related to the living indriids but adapted somewhat like baboons; and *Megaladapis*, a very large, vertically postured, short-limbed, and probably slow-moving climber with adaptive resemblances to the koala of Australia.

Tarsiers. The tiny *Tarsius*, which lives today in Southeast Asia, represents a link between the toothcombed strepsirhines and the anthropoids

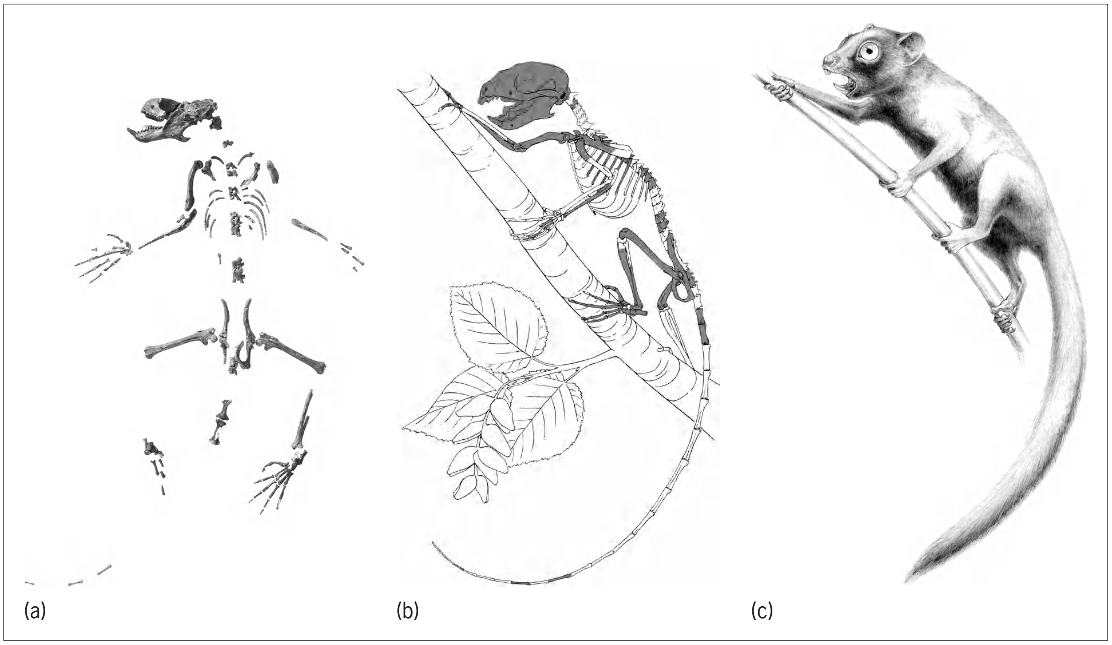


Fig. 2. ***Carpolestes simpsoni*, a 55.5-million-year-old plesiadapiform from Wyoming. (*a*) Skeleton as recovered, laid out in approximate anatomical position. (*b*) Reconstruction in life stance on tree branch, elements in dark gray are known. (*c*) External restoration. (*Courtesy of and © Jon Bloch and Doug Boyer, artwork by D. Boyer*)**

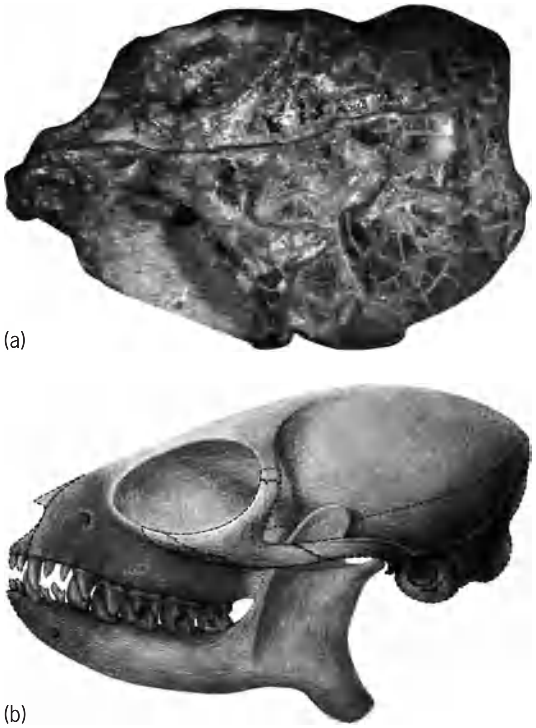


Fig. 3. Cranium of the omomyoid tarsiiform *Teilhardina asiatica*, from deposits dated ca 55 Ma in China. (*a*) Dorsal (top) view as found. (*b*) Left lateral reconstruction. (*Courtesy of and © Xijun Ni*)

(monkeys, apes, and humans). In many respects (such as lack of toothcomb, dry nasal area, molecular biology, and placentation), it is related to the anthropoids, but dentally it is usually thought primitive, although some authors have found similarities to some strepsirhines. The extinct Eocene omomyoids are close to tarsiers skeletally; they are often considered broadly ancestral to the anthropoids for that reason and also because some have monkeylike front teeth. Tarsierlike fossils have been recovered in Egypt and China from deposits of early Oligocene and middle Eocene age, respectively (30 and 45 Ma); in 2006, new material of the Chinese form supported a close link to living tarsiers. Another Chinese Eocene form, *Eosimias*, has been claimed to be an early anthropoid (higher primate), but some researchers argue that is also a tarsier relative, if more distant than the two preceding forms. Living tarsiers eat insects and small vertebrates, and are vertical clingers and leapers; their nocturnal habit is shown by their relatively huge eyes. They appear to live mainly in pairs, but they may form larger associations in some regions.

Higher primates. The anthropoids include three main groups of living animals and their extinct relatives; their divergence from a possibly omomyoid stock probably took place some 50 Ma. The monkeys of the New World and those of the Old World are of similar evolutionary grade, but the latter share a much more recent common ancestry with apes and humans, with which they are grouped in the infraorder Catarrhini.

The earliest known anthropoids may be represented by a few teeth (of *Algeripithecus*) from Algeria about 45 million years old, but better-preserved forms appear in the Fayum region of Egypt between 37 and 30 Ma following new finds and a revision of the geological dating in late 2005 and early 2006. Although now a desert, at the time this was a lush tropical forest through which meandered the sluggish proto-Nile river. Three main groups of early higher primates are represented: the oligopithecids, the parapithecids, and the propliopithecids. The former two appear to be only distantly related to any of the living monkeys or apes, although they may be termed monkeys in the broad sense, as their adaptations are similar to those of modern monkeys. The propliopithecids (species of *Propliopithecus*, including *Aegyptopithecus*) may be close to the common ancestry of later catarrhines (Old World anthropoids). These arboreal animals were the size of small cats, with apelike teeth, small brain, and limbs similar to those of the acrobatic South American atelines. Representatives of modern lineages begin to occur in the fossil record by about 28–20 Ma in both hemispheres. Most scholars subscribe to a colonization of South America by Fayum-age African anthropoids that crossed a then-narrower South Atlantic ocean on rafts of natural vegetation to reach the isolated island continent. *See* MONKEY.

New World primates. The platyrrhine or ateloid monkeys of South and Central America are divided into two families, of which the contents vary according to author. Here, the small marmosets and the common squirrel and capuchin monkeys are grouped into the family Cebidae, while the generally large-bodied spider-howler and saki-uakari groups are linked to the smaller titis and (probably) owl monkeys in the Atelidae. These two families differ in the relative robustness of their jaws and reduction of last molars. All living (and well-known extinct)

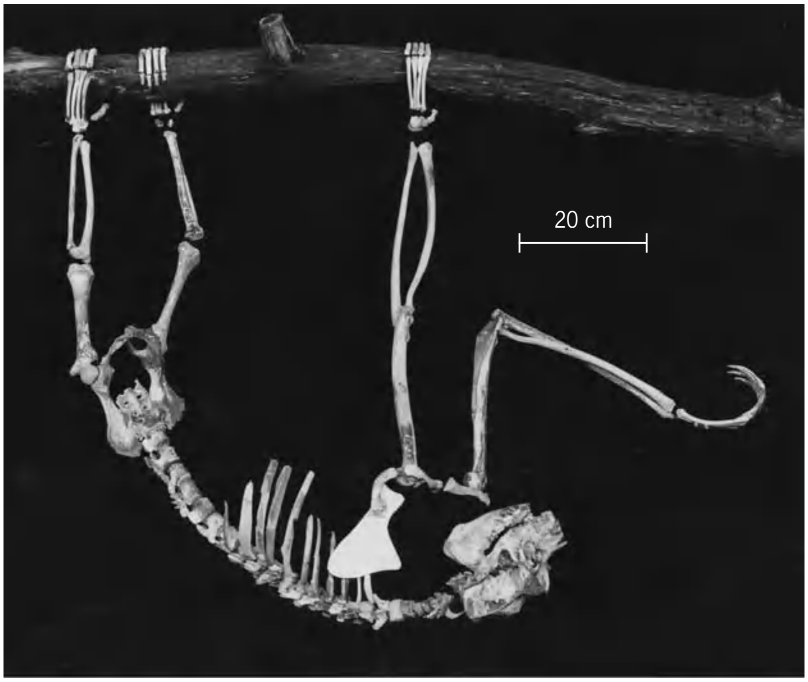


Fig. 4. Anjohibe specimen of *Palaeopropithecus ingens*, mounted in probable habitat posture. (*Photo courtesy of American Geological Institute*)

forms are arboreal and occupy forested areas between Argentina and southern Mexico. A distinctive evolutionary pattern observed in this group is the antiquity of the extant lineages as reflected by the close relationships of most of the few known fossils to modern genera.

The earliest fossil platyrrhine, 26-million-year-old *Branisella* from Bolivia, is as yet known only by teeth and jaw fragments. A probable size reduction of its last molars may suggest an evolutionary relationship to the Cebidae. It has also been suggested that *Branisella* might have lived in a savannah-like environment and been adapted to life at least partly on the ground, as opposed to all later platyrrhines. Two slightly later (about 20 Ma) species from Patagonia and Chile are also probable cebids, close to the ancestry of the living *Saimiri*, while a new species described early in 2006 is younger (16.4 Ma) and shows precocious brain size enlargement in the cebine lineage. Several forms dated 18–17 Ma represent early atelids, close to the living pitheciines (the saki-uakari and titi-owl monkey groups).

The largest number of fossil platyrrhines comes from the La Venta beds of Colombia, dated about 13 Ma. These include one or two species in each subfamily of Cebidae (Callitrichinae and Cebinae), one relative each of owl and howler monkeys, and four early members of the saki-uakari group. The last 100,000 years saw another flowering of extinct platyrrhine lineages. A cave site in eastern Brazil has yielded partial skeletons of "giant" relatives of the howler and spider monkeys, while several localities in the Caribbean produced controversial fossils perhaps related to howlers, saki-uakaris, and squirrel monkeys (or possibly representing a distinct lineage whose members came to resemble those other groups). *See* MONKEY.

Old World monkeys. The living Cercopithecidae are divided into two subfamilies, Colobinae and Cercopithecinae. The oldest cercopithecids are found in Africa, with a few fossil forms such as *Victoriapithecus* of 20–15 Ma probably predating the divergence of the modern subfamilies. The cercopithecines include a wide variety of forms, all of which share cheek pouches for temporary food storage and usually large incisors reflecting a fruit diet; colobines, by contrast, are more restricted in morphology, range, and behavior pattern, and all are leaf eaters with a complicated digestive tract to facilitate the low-nutrition diet. All cercopithecids possess ischial callosities (rump-sitting pads), which have been lost to a greater or lesser extent in apes and humans.

The earliest members of the two living subfamilies also are mainly African. One colobine jaw is known by 9 Ma, and from 7 to 5 Ma species of both cercopithecines and colobines become more abundant. Large collections of Old World monkey fossils have been recovered from East and South African sites (often in association with early human remains) in the 4–1.5 Ma interval. Cercopithecines are represented first by *Parapapio*, which was a semiterrestrial animal probably closer to the common ancestor of later forms, and then by several extinct species of the highly terrestrial living *Theropithecus*. At least three types of large colobine monkeys were also common, ranging from arboreal to terrestrial in adaptation. Many of these species were significantly larger than their living relatives: The biggest colobine may have weighed close to 50 kg (110 lb) [**Fig. 5**] while a *Theropithecus* dated about 1 Ma has been estimated nearly twice that large. Some of these large and slow-moving species may have been hunted and perhaps driven to extinction by early humans.

Fig. 5. Partly reconstructed skeleton (casts) of *Paracolobus chemeroni*, 3.2-million-year-old colobine monkey from Kenya; on exhibit at American Museum of Natural History, New York. (*Courtesy of and © Eric Delson*)

Cercopithecids entered Eurasia from Africa. *Mesopithecus pentelicus*, an 11–8-million-year-old colobine known in a geographical range from Germany through Afghanistan, is the best-represented Eurasian fossil monkey, with dozens of individuals recovered from sites in Greece and Bulgaria. It was semiterrestrial, perhaps like the living hanuman langur *Semnopithecus entellus*, sleeping in trees and feeding on the ground along watercourses. A possible descendant is *Dolichopithecus*, which lived in Europe from about 4.5 to 2.5 Ma and perhaps in Japan, Mongolia, and Siberia as well. It was even more terrestrially adapted, more so than any living colobine. The living macaques (*Macaca*) are widespread across eastern Asia and in North Africa, and their fossil record adds to that large range. Scattered specimens are known from North Africa after 7 Ma, and populations have been recovered across Europe from 5.5 Ma to about 100,000 years ago. Large and highly terrestrial relatives, such as *Paradolichopithecus*, inhabited Europe, central Asia, and eastern Asia from about 3 to 1.5 Ma.

Hominoids. The most humanlike of all primates are the apes, which form a group distinguished by generally large body size, relatively large brain, lack of an external tail, and advanced placentation pattern. Living forms include the lesser apes, or gibbons (*Hylobates*), placed in their own family Hylobatidae, and the several great apes: orangutan (*Pongo*), chimpanzee (*Pan*), and gorilla (*Gorilla*). The great apes and humans, along with some extinct relatives, are grouped as the Hominidae by some authors (as here), while others place only humans in the Hominidae and class all great apes in the Pongidae.

One of the earliest probable members of Hominoidea is *Proconsul*, of the East African Miocene, 23–14 million years old; a few teeth of a similar form date to 26 Ma. *Proconsul* is well known by most of its skeleton (**Fig. 6**). Several species ranged in size from small chimpanzee to small gorilla, with a somewhat chimplike skull, large projecting canine teeth, and limb bones seemingly adapted to quadrupedal running. However, *Proconsul* has few of the defining features of the ape group. It lacks the mobile shoulder characteristic of all living species, and there is some controversy over its tail: some workers believe that tail bones have been found but not always recognized as such. Thus, several researchers have removed *Proconsul* from the Hominoidea and placed it and several relatives as an advanced archaic catarrhine. For the present, *Proconsul* is retained as a hominoid belonging to a distinct archaic family of its own. Two other groups of roughly contemporaneous species (the Eurasian Pliopithecidae and the African "*Dendropithecus-group*") are clearly more primitive than *Proconsul*, although they have at times wrongly been included in Hominoidea, usually as purported relatives of the gibbons. The oldest fossil gibbons date only to about 1 Ma; presumably, their ancestors lived after *Proconsul* and entered Eurasia about the same times as did early Hominidae.

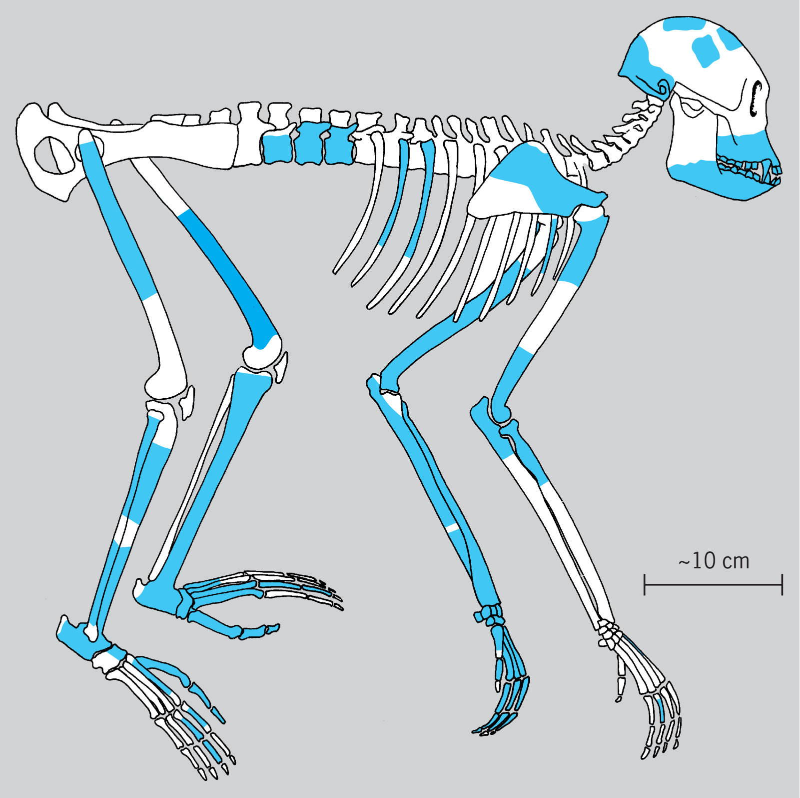


Fig. 6. Reconstructed skeleton of Proconsul africanus, an early hominoid (23-14 Ma) from East Africa. Known parts are shaded. (*After R. L. Ciochon and R. S. Corruccini, eds., New Interpretations of Ape and Human Evolution, Plenum Press, 1983*)

Several fossil apes from Africa dating to 20–14 Ma are often placed in the Hominidae because they appear to share "advanced" dental features with the modern great apes. But most of these species have recently been questioned as hominids because they lack apelike features of the postcranial skeleton (in the elbow, shoulder, hip, and knee joints, especially). Surprisingly, the oldest of these, *Morotopithecus*, may be the most modern, although only fragmentary postcranial elements are known. The youngest, *Kenyapithecus*, is similar to the first Eurasian hominoid, *Griphopithecus*, found in Turkey, Germany, Slovakia, and Austria between 16 and 13 Ma. None of these fossils are specially related to any living great ape, and they are often placed in the hominid subfamily Kenyapithecinae or even outside Hominidae.

Although interpretations vary, there appear to be three groups of Eurasian hominids between 13 and 7 Ma. *Dryopithecus* is characteristic of the Dryopithecinae, which may include the common ancestors of all later great apes (and humans). This animal has been known since the 1850s, but only in the 1990s were partial crania and skeletons described. These show several similarities to living apes in general but not to any specific forms.

Spread of modern ape ancestors. Most scientists today agree that, of the great apes, the orangutan is evolutionarily farthest from humans. This view is based mainly on the results of numerous studies of proteins and DNA sequences, which reveal great similarity among the genes of gorillas, chimpanzees, and humans, with orangutans, gibbons, and monkeys successively farther from that close-knit group. As a result, orangutans and their extinct relatives are here placed in the subfamily Ponginae, while African apes, humans, and their relatives are included in Homininae. The living orangutans (*Pongo*) inhabit a small range in the deep forests of Borneo and Sumatra; less than 1 Ma, the same species was also present as far north as southern China. *See* APES.

The orangutan lineage is, however, the oldest well-documented one among all catarrhines. Fossils from Pakistan and India known as *Sivapithecus* show facial, palatal, and dental architectures clearly specialized in the orang direction as far back as 13–12 Ma. Although slightly younger in time, *Ankarapithecus* from Turkey at approximately 10 Ma appears intermediate between *Sivapithecus* and *Dryopithecus* in facial morphology. These three genera seem to form a transformation series representing the approximate manner in which orang ancestry evolved from dryopithecines, and both *Ankarapithecus* and *Sivapithecus* are usually included in Ponginae. But whereas the living great apes (and human ancestors) share a complex of arm-bone features related to suspensory behavior and forelimb flexibility, the upper arm of *Sivapithecus* lacks at least one part of this complex. Thus, it is possible either that this lack denies *Sivapithecus* close relationship with orangs, or that Asian and African great apes independently evolved these features long thought to document their shared ancestry. Alternatively, the lack may merely represent a secondary specialization of *Sivapithecus* which happens to be similar to what is seen in more archaic apes such as the kenyapithecines.

This last option makes the most sense, as *Sivapithecus* in other ways appears to be somewhat more adapted to quadrupedal life on the ground than is the living orangutan.

Two larger species which probably belong to the Ponginae are placed in the genus *Gigantopithecus*: One dates to about 9–6 Ma from India and Pakistan, the other lived about 1.5–0.5 Ma in China and perhaps Vietnam. Hundreds of specimens, mostly isolated teeth, are known from China, and these document a species which was probably the largest primate that ever lived (perhaps weighing 200–400 kg or 440–880 lb, although teeth are not the best body part from which to estimate weight). *Gigantopithecus* was once thought to be near the ancestry of humans, but today it is usually considered a collateral descendant of *Sivapithecus*.

The pongines probably evolved in Asia from an arboreal dryopithecine (or even kenyapithecine) ancestry which expanded into moist forested environments. Such a habitat would have provided an abundance of fruit plus gritty and tough food objects (like fibrous and thick-skinned fruits or nuts), to which this group's dentition appears adapted. It was at first thought that these creatures might have spent significantly more time walking and feeding on the ground than modern or earlier apes, but partial limb bones suggest overall similarity of structure, and thus probably of behavioral function, to living arboreal and partly ground-living apes. They share with orangutans and some early humans a complex of dental-related features, including a thick molar enamel covering (to prolong tooth life with heavy wear), large molars in relation to estimated body size, strongly buttressed jaws (to take up stresses of heavy chewing), and some reorientation of the front teeth. A major factor in human dental evolution was facial shortening combined with reduction of the canine teeth and their transformation into incisorlike elements, but no Miocene fossils yet known show this feature clearly.

The origin of the Homininae is more problematic. The fossil ape *Graecopithecus* (also termed *Ouranopithecus*) is known from several Greek localities estimated to date between 9.6 and 8.7 Ma. Well-preserved facial material of this animal and of *Dryopithecus* recovered or reanalyzed in the 1990s has led different workers to suggest that one or both forms may lie near the split between Ponginae and Homininae or already on the hominine lineage, effectively close to the common ancestor of African apes and humans. This would agree well with the date of about 13–12 Ma for the earliest pongines on the assumption that both branches of the split would have developed their derived features contemporaneously. *Graecopithecus* shows more evidence of the downwardly bent face typical of African apes (a pattern termed klinorhynchy) and also very thick molar enamel (as expected in an early hominine), while *Dryopithecus* has thin enamel. Some of the differences between them may just reflect the far larger size of *Graecopithecus*, but at present this genus appears more derived in the direction of later hominines. At about the same time in Africa, the only known ape fossil is a single upper jaw which was named *Samburupithecus* in 1997 and which may also represent an early, gorillalike member of Homininae. These fossils imply that the hominines may have evolved in Eurasia and then returned to Africa after about 10 Ma. Previous workers often thought that the hominine lineage could be traced purely within Africa, but that hypothesis now appears less likely. The ancestry of the African apes is still a mystery: until mid-2005 no fossils had been found which clearly represent their lineage before or after separation from humans. However, a few teeth have now been reported from a Kenyan archeological site which appear to represent a form of chimpanzee, but so far these say little about the history of African ape adaptations.

Human evolution. During the last decade, three new African genera have been described that have been claimed to throw new light on the origin of the Hominini, the lineage including living humans but postdating the last common ancestor of humans and apes. Potentially the oldest of these is *Sahelanthropus*, found in 2002 in Chad, and dated in the 7–6 Ma range. Slightly younger is *Orrorin*, from Kenya, at about 6 Ma; and the youngest is *Ardipithecus*, reported from sites in Ethiopia dating between 5.8 and 4.3 Ma. Each of these is known from essentially different skeletal elements, and each has been disputed in one way or another, reflecting the fact that we do not as yet have an adequate conception of what the earliest hominins should be expected to look like. All, however, have been claimed, on the basis of tenuous evidence, to have been upright bipeds.

Several species of the genus *Australopithecus* are known from sites dating between 4.2 and 2.3 Ma in South Africa, Chad, and the Rift Valley from Tanzania through Ethiopia. They have crania with internal volumes ("brain size") of 400–500 ml, in the range of living gorillas; body weight, however, is far smaller, about 30 kg for females, up to 60 kg for males. The jaw musculature was small enough not to require strong crests on the outer surface of the skull. The dentition had moderately large incisors, relatively small and partly incisivized canines, thick-enameled molars, and a broadly "humanlike" arcade shape. The postcranial skeleton was clearly adapted to upright bipedalism (also indicated by footprint trails), although it differed in detail from modern humans, with relatively short legs. The diet of these forms was probably omnivorous, including fruits, seeds, and perhaps scavenged animal protein; there is no definitive evidence for toolmaking. *See* AUSTRALOPITHECINE; FOSSIL HUMANS.

The "robust" australopiths, genus *Paranthropus*, were generally younger (about 2.6–1.4 Ma) and more derived craniofacially. Their postcanine teeth were greatly enlarged and adapted to crushing and grinding hard food items, while the canines and incisors were strongly reduced except in the earliest species, *P. aethiopicus*. Sagittal and nuchal crests served as anchors for the large temporal musculature covering a vault enclosing an only slightly larger brain than in *Australopithecus*; the postcranial skeleton is poorly known, but body structure and size were probably not much different from that of its predecessor.

The earliest toolmakers probably belonged to species included in the living genus *Homo*, first appearing 2.5–2.2 Ma in Malawi, Kenya, and Ethiopia, soon after the oldest stone tools. There is some debate over the number of early *Homo* species, but it seems likely that at least three were contemporaneous if not actual neighbors in eastern Africa around 2 Ma. Soon afterward, *Homo erectus* (or according to some authors, a closely related species named *H. ergaster*) emigrated from Africa to eastern Eurasia where *H. erectus* persisted until at least 300 Ka (thousand years ago). Continental Europe may not have been colonized until about 1 Ma; almost all younger populations there appear to be part of a clade that included the Neanderthals, known from 200 to 30 Ka between Spain and central Asia. African populations of the last half million years appear to have evolved less directly into anatomically modern humans (*Homo sapiens*), some of whom dispersed via southwestern Asia to colonize the rest of the Earth after 150 Ka. *See* FOSSIL HUMANS; MAMMALIA.

Eric Delson; Ian Tattersall

Bibliography. E. Delson et al. (eds.), *Encyclopedia of Human Evolution and Prehistory*, 2d ed., 2000; J. G. Fleagle, *Primate Adaptation and Evolution*, 2d ed., 1998; W. C. Hartwig, *The Primate Fossil Record*, 2002; S. V. Nelson, *The Extinction of Sivapithecus*, 2003; J. H. Schwartz and I. Tattersall, *The Human Fossil Record*, vols. 1–4, 2001–2005; F. S. Szalay and E. Delson, *Evolutionary History of the Primates*, 1979.

Fossil seeds and fruits

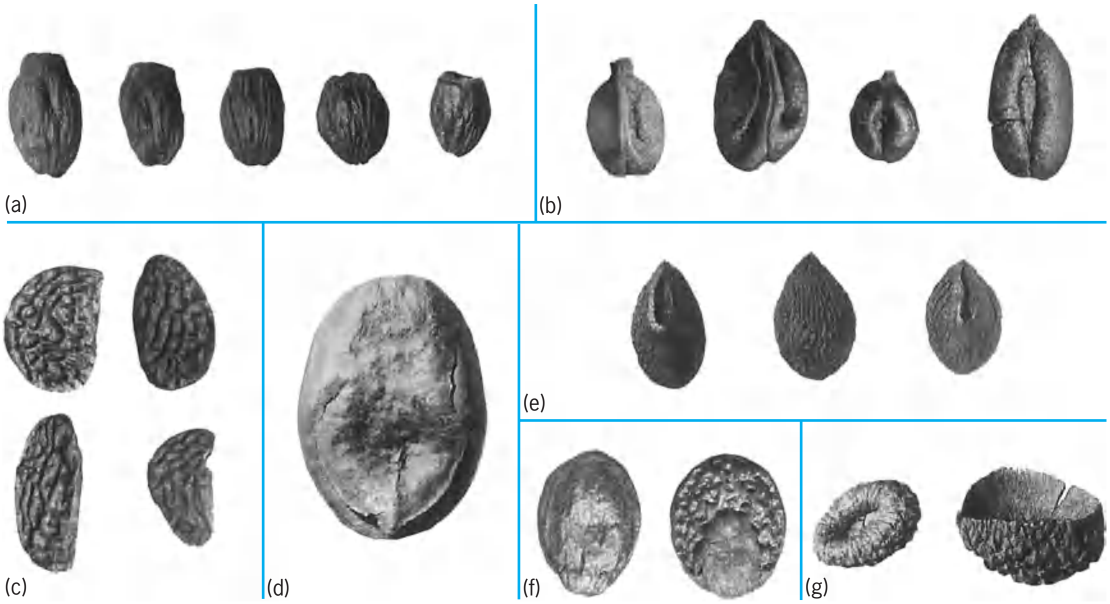
Seeds, ovules containing a fertilized egg and ready to be shed from the plant, are reproductive organs characteristic of both gymnospermous and angiospermous plants. In angiosperms (Magnoliophyta) an additional structure, the matured ovary, encloses one or more seeds to form a fruit. *See* MAGNOLIOPHYTA.

Seeds and fruits are less commonly found as fossils than are vegetative remains. They may be preserved structurally as casts, or as compressions which are sometimes found with leaf compressions. Seeds and fruits often occur in lignites. *See* LIGNITE.

The oldest known seed plants are of Mississippian age. Carboniferous seed plants include the extinct Cordaitales, probable conifer ancestors, and Pteriodospermae, seed plants with fernlike foliage. Both groups had similar seeds, basically like those of modern cycads. No embryos have been found in Paleozoic seeds. *See* CORDAITALES; PTERIDOSPERMS.

During the Mesozoic Era, all major modern groups of seed plants were represented, along with members of the declining cordaitalean and pteridospermous stocks. Among the most completely known Mesozoic seeds are those of the cycadeoids, extinct cycad relatives. Their silicified, beehive-shaped trunks may include shoots bearing numerous small seeds between thick scales. The Mesozoic Caytoniales, whose small seeds were borne in fleshy enclosures, were thought at one time to be ancestral angiosperms. Now they are recognized as gymnospermous forms, probably related to the pteridosperms. *See* CAYTONIALES; CYCADEOIDALES.

Angiosperm fruits are rare in Lower Cretaceous beds; Upper Cretaceous fruits are known from northern Africa, Long Island, New York, and elsewhere. Tertiary fruits and seeds have been found in numbers in the United States in the Brandon lignite of Vermont (see **illus.**) and in the Clarno Formation of central Oregon. Silicified cones of araucarian conifers are known from Patagonia. The best-known European Tertiary fruits and seeds are from the brown coals of



Early Tertiary fruits and seeds which were found in Brandon lignite of Vermont. (*a*) Endocarps of *Symplocos*. (*b*) Seeds of grape (*Vitis*). (*c*) Seeds of *Rubus*. (*d*) Endocarp of extinct species of *Nyssa*. (*e*) Seeds of fossil genus of family Rutaceae, which are related to *Phellodendron*. (*f*) Endocarps of a species of *Alangium*. (*g*) Acorn cups, *Quercus* sp. (*Courtesy of Elso S. Barghoorn*)

Germany and the Eocene London Clay Formation of England.

Important paleobotanical findings resulting from the study of fossil seeds and fruits include the knowledge obtained of the independent evolution of the seed habit in unrelated groups; the discovery that much Carboniferous fernlike foliage was borne on seed plants rather than on ferns; and the discovery that *Glossopteris*, an important plant in widespread Permian floras of the Southern Hemisphere, was a seed plant. Pyritized fruits from the London Clay Formation reveal the presence of many extinct genera along with modern genera in early Tertiary time. Morphological changes in herbaceous angiosperm seeds from sequences of Tertiary beds furnish data on rates of evolution in plants. Because plant classification is based primarily upon reproductive structures, fossil seeds and fruits provide highly reliable evidence for identification and interpretation of fossil plants. *See* PALEOBOTANY. **Richard A. Scott**

Bibliography. H. N. Andrews, *Studies in Paleobotany*, 1961; E. M. Reid and M. E. J. Chandler, *The London Clay Flora*, 1933; R. A. Scott, Fossil fruits and seeds from the Eocene Clarno Formation of Oregon, *Palaeontographica,* 96(B):66–97, 1954; A. C. Shukla and S. P. Misra, *Essentials of Paleobotany*, 2d ed., 1982; T. N. Taylor, *Paleobotany*: *An Introduction*, 1981.

Foucault pendulum

A pendulum or swinging weight, supported by a long wire, by which J. B. L. Foucault demonstrated in 1851 the rotation of Earth on its axis. Foucault used a 62-lb (28-kg) iron ball suspended on about a 200-ft (60-m) wire in the Pantheon in Paris. The upper support of the wire restrains the wire only in the vertical direction. The bob is set swinging along a meridian in pure translation (no lateral or circular motion). In the Northern Hemisphere the plane of swing appears to turn clockwise; in the Southern Hemisphere it appears to turn counterclockwise, the rate being 15 degrees times the sine of the local latitude per sidereal hour. Thus, at the Equator the plane of swing is carried around by Earth and the pendulum shows no apparent rotation; at either pole the plane of swing remains fixed in space while Earth completes one rotation each sidereal day. *See* DAY; INERTIAL GUIDANCE SYSTEM; PENDULUM; SCHULER PENDULUM. **Frank H. Rockett**

Bibliography. G. O. Abell, D. Morrison, and S. C. Wolff, *Exploration of the Universe*, 6th ed., 1991; C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, Berkeley Physics Course, vol. 1, 2d ed., 1973; K. R. Symon, *Mechanics*, 3d ed., 1971.

Foundations

Structures or other constructed works are supported on the earth by foundations. The word "foundation" may mean the earth itself, something placed in or on the earth to provide support, or a combination of the earth and the elements placed on it. The foundation for a multistory office building could be a combination of concrete footings and the soil or rock on which the footings are supported. The foundation for an earth-fill dam would be the natural soil or rock on which the dam is placed. Concrete footings or piles and pile caps are often referred to as foundations without including the soil or rock on which or in which they are placed. The installed elements and the natural soil or rock of the earth form a foundation system; the soil and rock provide the ultimate support of the system. Foundations that are installed may be either soil-bearing or rock-bearing. The reactions of the soil or rock to the imposed loads generally determine how well the foundation system functions. In designing the installed portions, the designer must determine the safe pressure which can be used on the soil or rock and the amount of total settlement and differential settlement which the structure can withstand. For a discussion of soil as a part of the foundation system *see* SOIL MECHANICS.

The installed parts of the foundation system may be footings, mat foundations, slab foundations, and caissons or piles, all of which are used to transfer load from a superstructure into the earth. These parts, which transmit load from the superstructure to the earth, are called the substructure (**Fig. 1**).

Footings. Footings or spread foundations are used to spread the loads from columns or walls to the

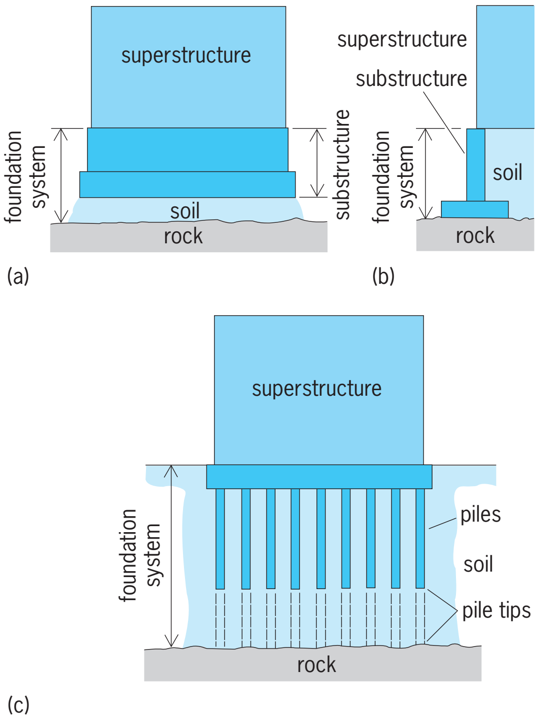


Fig. 1. Examples of foundation systems. (*a*) Structure supported on a foundation bearing on soil. (*b*) Structure supported on a foundation bearing on rock. (*c*) Structure supported by a pile foundation. The piles may be installed so that the pile tips terminate in soil and all support is derived from the soil, or they may be installed so that the tips penetrate to rock.

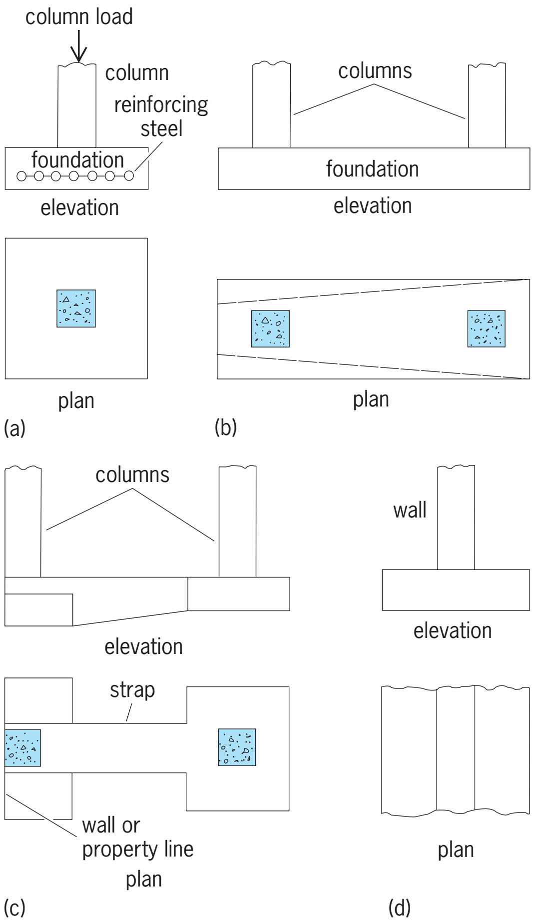


Fig. 2. Types of footings. (*a*) Column footing. (*b*) Combined footing, which can be either rectangular or trapezoidal. (*c*) Cantilevered footing. (*d*) Wall footing.

underlying soil or rock. Normally, footings are constructed of reinforced concrete; however, under some circumstances they may be constructed of plain concrete or masonry. When each footing supports only one column, it is square. Footings supporting two columns are called combined footings and may be either rectangular or trapezoidal. Cantilever footings are used to carry loads from two columns, with one column and one end of the footing placed against a building line or exterior wall. Footings supporting walls are continuous footings (**Fig. 2**).

The sizes of footings are determined by dividing the loads to be imposed at the base of the footings by the allowable bearing pressure which can be imposed on the soil or rock of the earth. Most building codes and textbooks on foundations contain tables listing allowable bearing pressures for various types of soil and rock; however, these tables give only general classifications and descriptions of the soil or rock and must be used with caution. More specific information about the soil or rock is normally obtained by drilling test borings, extracting soil or rock samples, performing laboratory tests on the samples, and making engineering analyses to determine suitable bearing pressures. In addition to bearing pressure, consideration must be given to the amount of settlement which may occur and the capability of the structure to withstand such settlement. If settlement is a problem it may be necessary to use an alternate foundation type rather than footings or to enlarge the footings and decrease the bearing pressure.

Grade beams may be used between exterior column footings to support walls, with the beams transferring the weight of the walls to the column footings. Beams are also used between interior column footings to act as braces or ties or to support interior walls. Retaining walls are thoses walls subjected to horizontal earth pressures due to the retention of earth behind them. The foundations for these walls must have sufficient frictional resistance with the soil or rock on which they rest so that they will not slide when subjected to the horizontal earth pressure. In addition, retaining walls must be designed so they will not overturn (**Fig. 3**). In frost-susceptible areas, footings must be placed below the frost line.

Mat foundations. Mat or raft foundations are large, thick, and usually heavily reinforced concrete mats which transfer loads from a number of columns or columns and walls to the underlying soil or rock. Mats are also combined footings, but are much larger than a footing supporting two columns. They are continuous footings and are designed to transfer a relatively uniform pressure to the underlying soil or rock. Mats are rigid and will act as a bridge over discontinuities in the soil or rock on which they are founded. Mats founded several meters below the ground surface, when combined with external walls, are termed floating foundations (**Fig. 4**). The weight

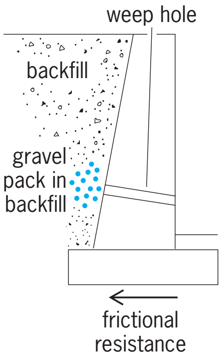


Fig. 3. Retaining wall. Gravel pack behind the wall prevents the weep hole from clogging with soil.

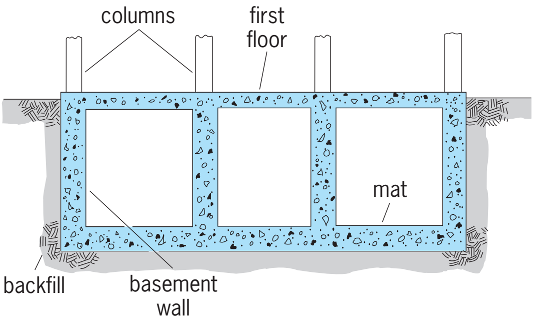


Fig. 4. Mat foundation, sometimes called a floating foundation, generally heavily reinforced in both directions.

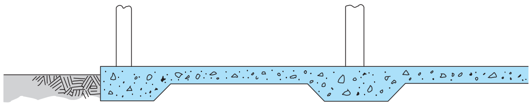


Fig. 5. Slab foundation.

of the soil excavated from the ground surface to the bottom of the mat may be equal to or approach the total weight of the structure. In this case, little or no new load is applied to the underlying supporting soil, and settlements of a structure may be minimal after construction.

Slab foundations. Slab foundations are used for light structures wherein the columns and walls are supported directly on the floor slab. The floor slab is thickened and more heavily reinforced at the places where the column and wall loads are imposed (**Fig. 5**).

Special problems. Ground water is a major problem in connection with the design and installation of foundations where a substructure is to be placed below the ground-water level. Well points, pumping from deep wells, or pumping from sumps are methods used to dewater construction sites during foundation installation. Other methods which are less often used are freezing, removal of water by electroosmosis, and the installation of cutoff walls made of piling or grout around the periphery from within the excavation. If dewatering operations are performed in an area surrounded by existing structures, precautions must be taken to protect them, as the lowering of the ground-water may cause the soil on which they are supported to subside.

If a basement is partially or totally below the ground-water level its walls must be designed to withstand the hydrostatic pressures of the water on the outside in addition to the pressure from the soil backfill. An alternate procedure is to install a permanent system to remove water outside the walls. Some substructures below ground-water level may at times be subjected to hydrostatic uplift forces which are greater than the downward forces imposed by the structures. In these cases, provisions must be made to anchor the structures to prevent them from floating upward.

Ground water also causes problems by infiltrating through basement walls, slabs, and joints into the basement itself. This can be prevented or reduced by providing an external permanent drainage syatem that carries water away from the basement, by encasing the walls and slabs in an inpermeable plastic membrane, or by coating the external walls with an asphaltic mastic to lower their permeability (**Fig. 6**). Combinations of the foregoing are also used. Retaining walls and abutments often have weep holes in the lower sections through which water accumulating behind the walls or abutments can escape. The water pressure behind the walls is relieved as the water flows through the walls into an open external drainage system.

Foundations placed on expansive soils are often subjected to distressing movements unless special precautions are taken. Expansive soils are those which swell and contract excessively with varying amounts of moisture. Problems can be overcome by installing foundations below the zone of significant change in moisture content and backfilling with non-expansive materials, by altering the soil with an admixture such as lime or cement so that volume changes do not occur, or by providing flexibility in the structure to accommodate movements.

Underpinning of foundations is often necessary, and it may either remedial or precautionary. Remedial underpinning is used to correct defects in existing foundations which may have settled excessively. If the structure is to be saved or returned to its original state, additional foundation support must be provided. Precautionary underpinning is used when new structures are to be installed adjacent to or beneath existing structures, as in the construction of city subways. Underpinning of foundations is a specialized construction technique. The work is generally performed in a confined space, such as the basement of a building, or in small pits excavated outside a building area. It is necessary to provide support for the loads of the existing structure while new foundations are installed. The new foundations may be footings which are placed deeper in the ground than the original foundations, or they may be piles or caissons.

Underpinning of a wall footing may be performed by excavating pits adjacent to and beneath existing foundations. The pits are small, some 3 ft (0.9 m) wide by 4 ft (1.2 m) long. Horizontal sheeting is placed in the pits as excavation proceeds to prevent caving of the walls and undermining of the structure being underpinned. When the new bearing stratum is reached, forms are placed in the pit, and concrete is poured from the new bearing stratum up to within 3 in. (76 mm) of the bottom of the old footing. After

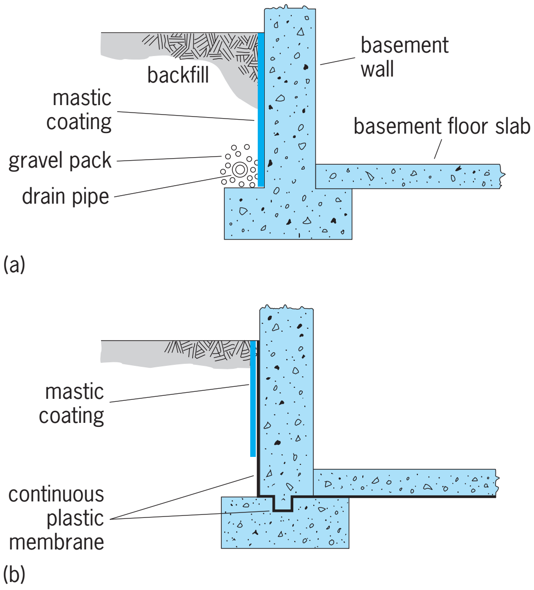


Fig. 6. Methods of reducing water infiltration into basements. (*a*) External drain pipe. (*b*) Walls, floor slab, and joints encased in a plastic membrane.

the new concrete has hardened, the 3-in. (76-mm) space is packed by hand with a mixture of sand, cement, and a small amount of water. Called grout, the mixture is packed very tightly into the space between the top of the new footing and the underside of the old footing. The pit underpinning process is repeated throughout the entire length of the wall footing. The resulting new foundation may be a continuous wall or intermittent piers. *See* BRIDGE; BUILDINGS; CAISSON FOUNDATION; DAM; PILE FOUNDATION; RETAINING WALL. **Gardner M. Reynolds**

Bibliography. R. Ahlvin and V. A. Smoots, *Construction Guide for Soils and Foundations*, 2d ed., 1988; J. E. Bowles, *Foundation Analysis and Design*, 5th ed., 1996; H. Y. Fang, *Foundation Engineering Handbook*, 2d ed., 1991; S. Hansbo, *Foundation Engineering*, 1994; L. ZeeVaert, *Foundation Engineering for Difficult Subsoil Conditions*, 2d ed., 1982.

Four-bar linkage

A basic linkage mechanism used in machinery and mechanical equipment. The term has been applied to three types of linkages: plane, spherical, and skew.

The plane four-bar linkage (**Fig. 1**) consists of four pin-connected links forming a closed loop, in which all pin axes are parallel. The spherical four-bar linkage (**Fig. 2**) consists of four pin-connected links forming a closed loop, in which all pin axes intersect at one point. The skew four-bar linkage (**Fig. 3**) consists of four jointed links forming a closed loop, in which crank 2 and link 4 are pin-connected to ground 1 and the axes of the pins are generally nonparallel and nonintersecting; coupler 3 is connected to crank 2 and link 4 by ball joints.

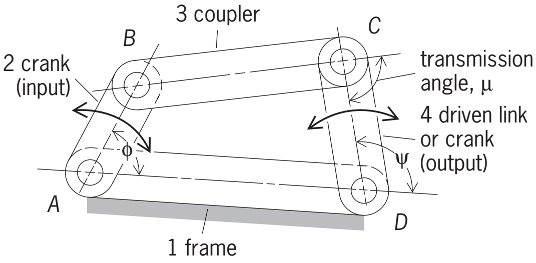


Fig. 1. Plane four-bar linkage.

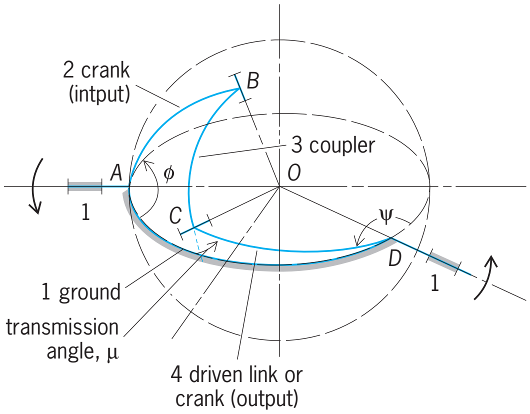


Fig. 2. Spherical four-bar linkage.

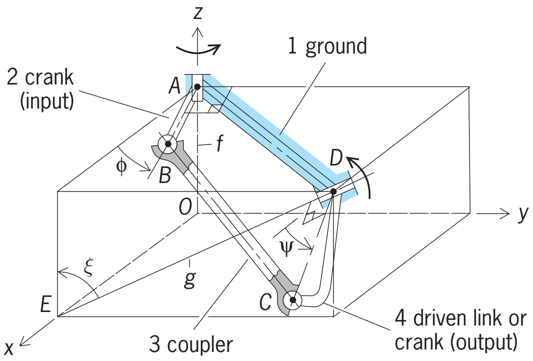


Fig. 3. Skew four-bar linkage. *OA* = *f*; *ED* = *g*; *OE* = common perpendicular between axes of pin joints at *A* and *D*.

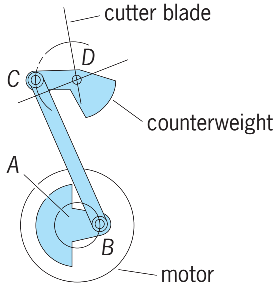


Fig. 4. Plane crank-and-rocker as used in an electric razor.

Mechanical function. Four-bar linkages are most frequently used to convert a uniform, continuous rotation (the motion of crank 2) into a nonuniform rotation or oscillation (the motion of link 4). Depending on whether the axes of the connected shafts (attached to links 2 and 4) are parallel, intersecting, or skew, a plane, spherical, or skew four-bar linkage is used. In instrument applications the primary function of the linkage is the conversion of motion, while in power applications both motion conversion and power transmission are fundamental. Four-bar linkages also are used to impart a prescribed motion to the coupler, as well as for other purposes.

Types of motions. Each of the above linkages can be proportioned for three types of motion, or linkage types: crank-and-rocker, drag, and double-rocker.

Crank-and-rocker linkages. In this motion the crank (link 2) is capable of unlimited rotation, while the output link (link 4) oscillates or rocks through a fraction of one turn (usually less than 90°). This is the most common form of the plane and the skew four-bar linkage, and is used in machinery and appliances of all types, for example, in an electric razor (**Fig. 4**). The most common spherical crank-and-rocker motion is known as the wobble plate. In wobble-plate motion the angle between the pin axes on the crank is less than 90°, while for each of the other links the angle between the pin axes is equal to 90°. The wobble motion is that of the coupler. It has been used in axial engines and elsewhere, for example, in dough kneading (**Fig. 5**), in which the dough is stirred by

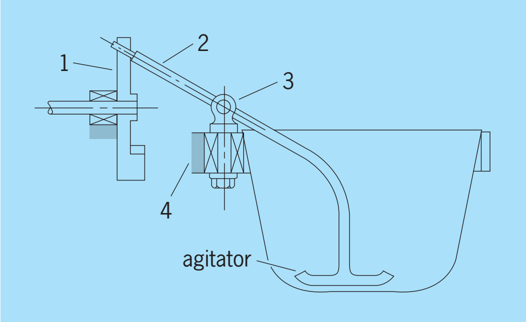


Fig. 5. Dough-kneading mechanism.

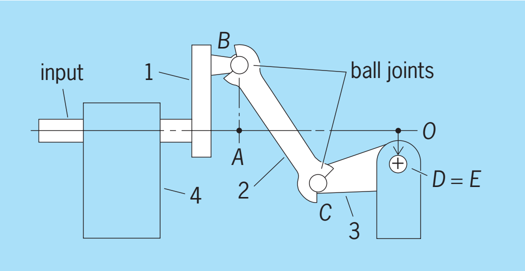


Fig. 6. Skew crank-and-rocker linkage.

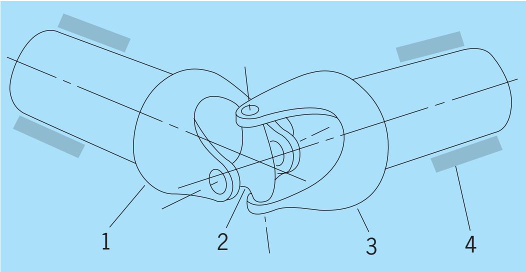


Fig. 7. Hooke joint, a universal joint with spherical drag linkage.

an agitator integral with the coupler. A skew four-bar linkage of crank-and-rocker proportions is shown in **Fig. 6**; the angle between the axes of crank 2 and rocker 4 is a right angle. Such motions have been used in textile machinery, tobacco-processing machinery, and generally for "going around corners."

Drag linkages. In drag linkages the motions of cranks 2 and 4 are both capable of unlimited rotations. The plane drag linkage has been used for quick-return motions, but care must be taken to avoid mechanical interference of the cranks with the coupler. The skew drag linkage is rarely used. The most common drag linkage is the spherical drag linkage. One such linkage is the Hooke-type universal joint, or Hooke joint, in which the angle between the pin axes of each moving link is a right angle (**Fig. 7**). The Hooke joint is used to connect intersecting shafts, for example, in automotive drive lines. The angular-velocity ratio between the connected shafts depends on the shaft angle (which is usually less than or equal to 30°) and is not constant. *See* UNIVERSAL JOINT.

Double-rocker linkages. In such linkages (**Fig. 8**) neither crank 2 nor 4 is capable of complete rotations. Such motions occur in hand tools and mechanical equipment in which only limited rotations are required, such as the oscillating motion in certain household fans.

Grashof's inequality. In the case of plane four-bar linkages the proportions which determine the type of motion are governed by Grashof's inequality: the sum of the lengths of the longest and shortest links is less than the sum of the lengths of the two intermediate links. If the inequality is satisfied and either link 2 or link 4 is the shortest link, the linkage is a crank-and-rocker (the shortest link functioning as the crank). If the equality is satisfied and the frame is the shortest link, the linkage is a drag linkage. In all other cases the linkage is a double-rocker linkage or a folding linkage. A folding linkage occurs when Grashof's inequality becomes an equality. Such linkages include those having parallelogram proportions, such as occur in the parallel and antiparallel equal-crank linkages. The latter can be used only, however, if provision is made for a positive drive while passing through the folded position or positions.

For spherical four-bar linkages a simple extension of Grashof's inequality is available, but for skew four-bar linkages the criteria governing motion types are more involved.

Transmission characteristics. For the transfer of power and motion in crank-driven four-bar linkages, the linkages should be proportioned to avoid lockup positions. In such positions a turning effort applied to the crank will not produce a turning effort applied to the output link. In plane and spherical four-bar linkages, a lockup position occurs when the pin axes on the coupler and output link are coplanar. The transmission angle μ (Figs. 1 and 2) is then equal to 0° or 180°. For efficient force transmission in plane and spherical four-bar linkages the maximum deviation Δ of the transmission angle μ from 90° should be minimized (usually $\Delta \leq 30^\circ$, at least in high-speed operation). Analogous but more complicated conditions apply to skew four-bar linkages. The Hooke joint and wobble plate are well proportioned for force transmission, as are the "centric" plane crank-and-rocker linkages for which $(AB)^2 + (AD)^2 = (BC)^2 + (CD)^2$. The determination of crank-and-rocker proportions yielding optimum force transmission as a function of

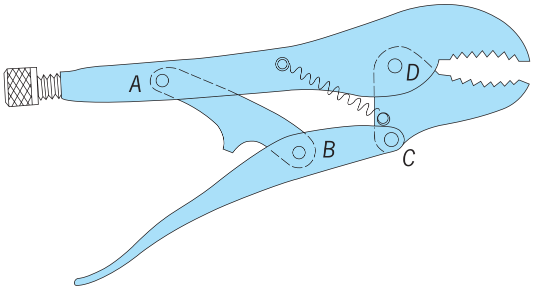


Fig. 8. Linkage-type pliers.

rocker swing angle and corresponding crank rotation has been solved as well.

Angular displacements. Equation the angular displacements of crank 2 and link 4 for the plane four-bar linkage and for a special case of the skew four-bar linkage—the "simply skewed" ($f = g = 0$; Fig. 3).

$$\begin{aligned} f(t,u) = \{&[(a-b-d)^2 - c^2] \\ &+ [(a+b-d)^2 - c^2]t^2\}u^2 \\ &- 8\lambda bdtu + [(a+b+d)^2 - c^2]t^2 \\ &+ [(a-b+d)^2 - c^2] = 0 \qquad (1) \end{aligned}$$

Equation (2) gives the displacements of crank 2 and link 4 for the spherical four-bar linkage.

$$\begin{aligned} f(t,u) = \{&[\cos(a-b-d) - \cos c] \\ &+ [\cos(a+b-d) - \cos c]t^2\}u^2 \\ &+ 4tu \sin b \sin d \\ &+ [\cos(a-b+d) - \cos c] = 0 \qquad (2) \end{aligned}$$

These equations are useful not only in determination of the angular displacements of links 2 and 4, but in the determination of their angular velocities and accelerations, which follow by differentiation.

Link sizes are expressed in ordinary length units, such as inches (for the plane and skew four-bar linkages), and in great circular arcs (in degrees, say, for the spherical four-bar linkage). Let *a, b, c, d* denote the lengths (linear of arc) of ground (*AD*), crank, coupler, and output link, respectively, and let ξ denote the angular offset between pin axes of links 2 and 4 of the simply skewed four-bar linkage. $\lambda = 1$ (in a plane four-bar linkage) or $\cos \xi$ (in a simply skewed four-bar linkage). If ϕ and ψ denote the angular positions of links 2 and 4, with $t = \tan \phi/2$ and $\mu = \tan \psi/2$, the displacement equations, $f(t,u) = 0$, giving the relationship between angular displacements of links 2 and 4, are as in Eqs. (1) and (2).

Design. This involves proportioning the linkages for given motion and power requirements and considering kinematics, dynamics, strength, lubrication, wear, balancing, tolerances, actuation, control and other factors. At low speeds the links can often be treated as rigid bodies, but at high speeds the elasticity needs to be included. The design procedures can be analytical, graphical, experimental, or a combination of these.

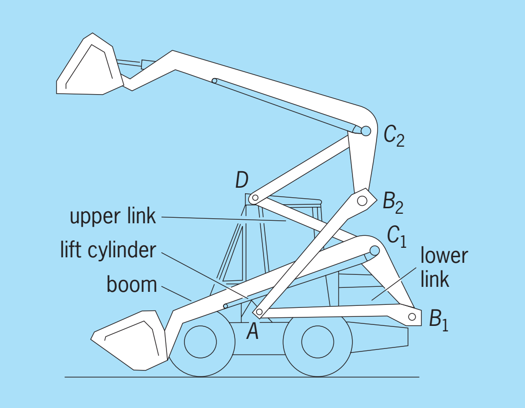


Fig. 9. Front loader with the bucket mounted on coupler *BC* of the four-bar linkage *ABCD*.

Due to the nonlinearity of the motion of four-bar linkages, their design is often time-consuming and difficult. However, the determination of linkage proportions for certain standard design requirements has been worked out in an efficient manner. These include transmission-angle optimization, the coordination of the rotations of links 2 and 4, and the generation of a particular coupler motion or point path (**Fig. 9**). In addition, general-purpose computer codes for the kinematic design and dynamic analysis of linkages, including four-bar linkages, have been developed.

Design trends emphasize high-speed operation and increasingly sophisticated motion and power-transmission requirements. *See* MECHANISM; SLIDER-CRANK MECHANISM; STRAIGHT-LINE MECHANISM.

Ferdinand Freudenstein

Bibliography. A. G. Erdman, *Modern Kinematics: Developments in the Last Forty Years*, 1993; A. G. Erdman and G. N. Sandor, *Mechanism Design*: *Analysis and Synthesis*, vol. 1, 3d ed., 1996; B. Paul, *Kinematics and Dynamics of Planar Machinery*, 1979; S. B. Tuttle, *Mechanisms for Engineering Design*, 1967.

Fourier series and transforms

Mathematical tools for the analysis of functions through decomposition into sinusoids. In 1822, J. Fourier proposed that ordinary mathematical functions could be represented by a sum of sinusoids, even though at first sight it might seem that the blandness and repetitive nature of the sinusoid ill suited it to accommodate the variety of functions in general. The advantage of such a representation is this: A differential equation that is difficult to solve under the untidy given external conditions that often arise in technology may very well be soluble under an external condition that is simply sinusoidal. If solutions can be obtained for each of the constituent sinusoids that make up the given external function, then perhaps the required solution will be the sum of the separate simple solutions. This was indeed so in the case studied by Fourier, who was interested in the differential equation governing the diffusion of heat in homogeneous solids; in fact, it is often the case that solving for sinusoids permits synthesis of the required solution. When this does happen, the solution is itself composed of sinusoids and there is said to be a sinusoidal response to sinusoidal input; the differential equation will also exhibit linearity combined with shift invariance. *See* DIFFERENTIAL EQUATION; HARMONIC ANALYZER; LINEARITY; TRIGONOMETRY.

Fourier series. A Fourier series is a sum of a constant and any number of sinusoidal functions, say of t, with the property that the frequencies of the component functions belong to an arithmetic sequence. An example is expression (1), where ω_0 is the fundamental frequency. The phase of each term is arbitrary, as is the amplitude, and in particular the amplitude associated with the fundamental term may be zero. Allowing an amplitude c_n for the nth term, and a phase α_n, the most general time-dependent Fourier series is given by expression (2). An equivalent expression, which uses two amplitudes a_n and b_n instead of one amplitude and a phase, is (3), where the coefficients in the two expressions are related by Eqs. (4). If the fundamental frequency is unity, the Fourier series in its simplest form results in expression (5).

$$2 + \cos(\omega_0 t + 0.4) + {}^{1}\!/_{2}\cos(2\omega_0 t + 0.5) + {}^{1}\!/_{4}\cos(3\omega_0 t - 0.2) \tag{1}$$

$$\sum_{n=0}^{\infty} c_n \cos(n\omega_0 t + \alpha_n) \tag{2}$$

$$\sum_{n=0}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t) \tag{3}$$

$$\begin{aligned} a_n &= c_n \cos \alpha_n \\ b_n &= -c_n \sin \alpha_n \end{aligned} \tag{4}$$

$$a_0 + (a_1 \cos t + b_1 \sin t) + (a_2 \cos 2t + b_2 \sin 2t) + \cdots \tag{5}$$

Fourier series have wide application in physics because systems that vibrate in response to a stimulus often do so periodically; a violin string offers an example. The reason that the vibrations of a clamped uniform string are closely periodic is that the velocity of low-amplitude transverse waves on a stretched uniform string is dependent, to a good approximation, on only the string tension and the string mass per unit length. *See* VIBRATION.

Not all natural vibrations are periodic; examples are furnished by drums, bells, bars, and horns, where the more general trigonometric series applies, namely expression (6). The frequencies ω_n in such cases are not necessarily in arithmetic sequence.

$$\sum_{n=0}^{\infty} c_n \cos(\omega_n t + \alpha_n) \tag{6}$$

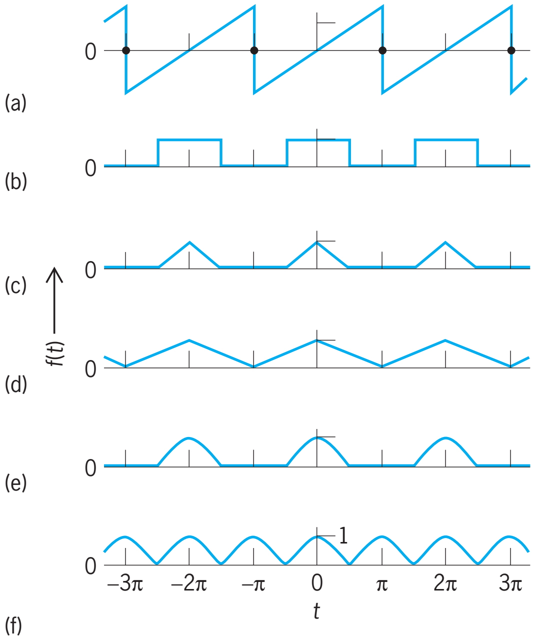
No term in the series (5) is altered in value if t is replaced by $t + 2\pi$; consequently if the series has a finite sum, that sum will be a function of t that is periodic in t with period 2π. In the discussion of Fourier series, attention may therefore be restricted to periodic functions. A real function $f(t)$ defined on the closed interval $[-\pi, \pi]$ and periodic with period 2π is said to have a Fourier series (5), where the coefficients are given by Eq. (7), and for $n > 0$, by Eqs. (8) and (9).

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt \tag{7}$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt \tag{8}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt \tag{9}$$

Existence conditions. When $f(t)$ is not suitably behaved, one or more of the integrals for the Fourier coefficients may not exist. However, the functions encountered in everyday experience are not apt to cause difficulty; even multiple discontinuities and poles are not sufficient to exclude a function. However, the existence conditions have been in contention since the time of Fourier and were responsible for major developments in the theory of functions of a real variable. The flavor of the discussion is given by noting that in the case of $f(t) = \cos(1/t)$, a function that has an infinite number of maxima in any finite interval around $t = 0$, no matter



Examples of (*a–c*) macrodasyid and (*d–f*) chaetonotid gastrotrichs. (*a*) *Xenodasys.* (*b*) *Dolichodasys.* (*c*) *Thaumastoderma.* (*d*) *Draculiciteria.* (*e*) *Halichaetonotus.* (*f*) *Aspidiophorus.* (*part a after C. Schoepfer-Sterrer, Chordodasys riedli gen. nov., spec. nov., a macrodasyoid gastrotrich with a chordoid organ, Cah. Biol. Mar., 10:391–404, 1969; b after G. D. Gagne, Dolichodasys elongatus n.g., n. sp., a new macrodasyid gastrotrich from New England, Trans. Amer. Microsc. Soc., 96:19–27, 1977; c after A. Remane, Klassen und Ordnungen des Tierreichs, vol. 4(2), Akad. Verlagsges., Leipzig, 1936; d after P. Luporini et al., Contribution à la connaissance des Gastrotriches des côtes de Toscane, Cah. Biol. Mar., 12:433–455, 1971; e after H. Mock, Chaetonotoidea (Gastrotricha) der Nordseeinsel Sylt, Mikrofauna Meeresbodens, 78:1–107, 1979; f after W. D. Hummon, Gastrotricha from Beaufort, North Carolina, U.S.A., Cah. Biol. Mar., 15:431–446, 1974*)

how narrow, the a_n do not exist. However, coefficients do exist for the function $t \cos(1/t)$, which also has an infinite number of maxima and minima. This situation can be covered by saying that the coefficients will exist if the arc length of $f(t)$ is finite. Neither of the above functions can be graphed, and the discussion belongs to a realm outside everyday experience. Furthermore, there are functions with infinite arc length that do have Fourier coefficients. Even if the coefficients exist, the sum of the Fourier series need not equal the original function. The rising importance of finite mathematics, largely associated with computers, means that, increasingly, users of Fourier series are concerned only with entities that can be represented by finite sets of data having finite values; the remainder of this article addresses such applications. *See* REAL VARIABLE; SERIES.

Examples. A simple periodic function is the periodic sawtooth waveform which equals ${}^1\!/_2 t$ in the interval $-\pi < t < \pi$. The coefficients a_n, including a_0, are all zero because of odd symmetry, and the values of b_n are easily obtained by integration to yield the series (10). This particular series converges to ${}^1\!/_2 t$ in the open interval $(-\pi, \pi)$ and had been summed by L. Euler before 1800. For other values of t (except odd multiples of π) it converges to the sawtooth generated by replication at multiples of 2π. Where $t = \pm\pi, \pm 3\pi, \pm 5\pi, \ldots$ the series also converges; the sum is zero. It is a general property that at a point of simple discontinuity, where the periodic function jumps, the series sums to the midpoint between the limiting values approached from left and right (**illus.** *a*).

$$\sin t - {}^1\!/_2 \sin 2t + {}^1\!/_3 \sin 3t - {}^1\!/_4 \sin 4t + \cdots \tag{10}$$

The Fourier series for some simple waveforms (see illus.) given in Eqs. (11). The sinc function is defined by $\operatorname{sinc} x = (\pi x)^{-1} \sin \pi x$.

Sawtooth:

$$a_0 = 0 \quad a_n = 0 \quad b_n = (-1)^{n-1} n^{-1} \tag{11a}$$

Rectangles:

$$a_0 = {}^1\!/_2 \quad b_n = 0 \quad a_n = \operatorname{sinc} {}^1\!/_2 n \tag{11b}$$

Spaced triangles:

$$a_0 = {}^1\!/_4 \quad b_n = 0 \quad a_n = {}^1\!/_2 \operatorname{sinc}^2 {}^1\!/_4 n \tag{11c}$$

Adjacent triangles:

$$a_0 = {}^1\!/_2 \quad b_n = 0 \quad a_n = \operatorname{sinc}^2 {}^1\!/_2 n \tag{11d}$$

Spaced cosinusoidal half-cycles:

$$\begin{aligned} &a_0 = 1/\pi \quad b_n = 0 \\ &a_n = {}^1\!/_2 \operatorname{sinc}\left({}^1\!/_2 n + {}^1\!/_2\right) + {}^1\!/_2 \operatorname{sinc}\left({}^1\!/_2 n + {}^1\!/_2\right) \end{aligned} \tag{11e}$$

Adjacent cosinusoidal half-cycles:

$$\begin{aligned} &a_0 = 2/\pi \quad b_n = 0 \quad a_1 = 0 \\ &a_n = -4 \cos\left({}^1\!/_2 n\pi\right)/[\pi(n^2 - 1)] \text{ for } n > 1 \end{aligned} \tag{11f}$$

Results for waveforms formed by replicating other pulse shapes can be obtained by integration; alternatively, the coefficients can be obtained from the Fourier transform of a single pulse, as explained below.

Fourier transforms. The integrals in Eqs. (8) and (9), which are Fourier integrals, have been of great historical importance. In modern times they have gained additional importance for Fourier analysis of aperiodic functions and the corresponding spectral analysis or diffraction of electromagnetic radiation (especially for chemical analysis by ultraviolet, visible, and infrared light, and by x-rays), microwaves, seismic waves, sound waves, ocean waves and tides, and vibrations. The advent of fast computing has also given importance to algorithms such as the fast Fourier and fast Hartley transforms, which have facilitated progress in plasma physics and fluid dynamics, where numerical multidimensional analysis was previously infeasible. *See* ACOUSTIC SIGNAL PROCESSING; DIFFRACTION; HOLOGRAPHY; INFRARED SPECTROSCOPY; INTERFEROMETRY; OCEAN WAVES; SEISMOLOGY; X-RAY CRYSTALLOGRAPHY.

A function $f(x)$ that is not periodic does not analyze into a discrete set of terms, each of which has finite energy carried by a single frequency, but rather into a spectral continuum. This continuum is analogous to the spectrum of sunlight, whose energy spectrum is describable in terms of the energy per hertz or other unit of frequency. A function $f(x)$ and its amplitude spectrum or Fourier transform $F(s)$ are related by Eqs. (12). If x represents time in seconds,

$$\begin{aligned} F(s) &= \int_{-\infty}^{\infty} f(x) e^{-i2\pi sx} dx \\ f(x) &= \int_{-\infty}^{\infty} F(s) e^{i2\pi sx} ds \end{aligned} \tag{12}$$

then s is frequency in cycles per second, or hertz; while if x represents distance, then s is spatial frequency in cycles per unit of x. In two dimensions the Fourier transform is given by Eq. (13), where u and v are spatial frequencies measured in the x and y directions respectively. *See* LAPLACE TRANSFORM.

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(ux+vy)} \, dx \, dy \tag{13}$$

A few special functions, defined in Eqs. (14), (15), and (16), are useful in specifying basic Fourier

$$\operatorname{rect} x = \begin{cases} = 1, & |x| < 0.5 \\ = 0, & |x| > 0.5 \end{cases} \tag{14}$$

$$\Lambda(x) = \begin{cases} 1 - |x|, & |x| < 1 \\ 0, & |x| > 1 \end{cases} \tag{15}$$

TABLE 1. Some Fourier transform pairs

$e^{-\pi x^2}$	$e^{-\pi s^2}$
sinc x	rect s
$\text{sinc}^2\, x$	$\Lambda(s)$
1	$\delta(s)$
$\cos \pi x$	${}^1\!/_2\delta(s + {}^1\!/_2) + {}^1\!/_2\delta(s - {}^1\!/_2)$
$\sin \pi x$	${}^1\!/_2 i\delta(s + {}^1\!/_2) - {}^1\!/_2 i\delta(s - {}^1\!/_2)$
$\text{III}(x)$	$\text{III}(s)$

$$\text{III}(x) = \sum_{n=-\infty}^{\infty} \delta(x - n) \tag{16}$$

transform pairs (**Table 1**). The unit impulse function, or Dirac delta function, is represented by $\delta(x)$, while the infinite train of unit impulses at unit spacing is represented by the shah function $\text{III}(x)$, an entity that is its own Fourier transform. Sample values of a function $f(x)$ defined on continuous x are expressible by the product $\text{III}(x)f(x)$, while a periodic function derived by replicating $f(x)$ is expressible by the convolution $\text{III}(x) * f(x)$. The asterisk notation for convolution is defined by Eq. (17).

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(x - u)g(u)\, du \tag{17}$$

Generation of Fourier pairs by theorem. A number of general theorems for the Fourier transform (**Table 2**) can be used to generate numerous additional Fourier transform pairs. For example, knowing that $f(x) = \exp(-\pi x^2)$ has Fourier transform $F(s) = \exp(-\pi s^2)$, it is not necessary to resort to further integration to find the Fourier transform of, say, $\exp(-x^2)$. The result can be deduced from the similarity theorem, which states that $f(ax)$ has Fourier transform $|a|^{-1}F(s/a)$. Hence $\exp(-x^2)$ has Fourier transform $\pi^{1/2} \exp(-\pi^2 s^2)$.

Periodic functions. A periodic function $p(x)$ of period T can be composed by replicating a given function $f(x)$ that is zero where $|x| > T/2$. Then $p(x)$ can be expressed as the convolution of $f(x)$ with a train of unit impulses at spacing T, namely $T^{-1}\text{III}(x/T)$. The factor T^{-1} is required because an impulse when stretched becomes stronger, a property that is expressed by $\delta(x/a) = |a|\ \delta(x)$. Thus, $p(x)$ is given by Eq. (18).

$$p(x) = T^{-1}\text{III}(x/T) * f(x) \tag{18}$$

Application of the convolution theorem yields the Fourier transform $P(s)$ as a simple sampling product, Eq. (19), which is equal to zero except where

$$P(s) = \text{III}(Ts)F(s) \tag{19}$$

$Ts = n$ (an integer). The only information retained about $F(s)$ is the set of sample values, $F(n/T)$, $-\infty < n < \infty$. These values must be equivalent to the Fourier coefficients; thus $a_0 = T^{-1}$ and, for $n > 0$, the relationship is given by Eq. (20).

$$a_n - ib_n = 2T^{-1}F(n/T) \tag{20}$$

Discrete Fourier transform. In numerical practice, the equivalent of a given function $f(x)$ of the continuous independent variable x is a finite number N of data values associated with a serial number τ, customarily taken as running from 0 to $N - 1$. Then there is a discrete Fourier transform $F(\nu)$ defined by Eq. (21). The quantity ν is an integer running from 0

$$F(\nu) = N^{-1} \sum_{\tau=0}^{N-1} f(\tau) e^{-i2\pi(\nu/N)\tau} \tag{21}$$

to $N - 1$. The factor N^{-1} ensures that $F(0)$ is the mean value of the N data values and is thus analogous to a_0, the Fourier coefficient that represents the mean value over one period. For $1 \le \nu \le N/2$, the quantity ν/N is analogous to frequency as measured in cycles per spacing of τ values. Values of ν exceeding $N/2$ represent negative frequencies $(\nu - N)/N$. The lowest positive frequency, or fundamental, is $1/N$, and the highest frequency is ${}^1\!/_2$. The shortest period representable by data is necessarily two units of τ, since at least two samples per period are required. The purpose of computing the discrete transform may be to present the power spectrum $|F(\nu)|^2$. In other cases the purpose may be to manipulate the spectrum by removing low frequencies, enhancing high frequencies, or removing high frequencies representing unwanted noise. Having done this, one returns to the data domain. Given a discrete transform $F(s)$, the corresponding data-domain function $f(\tau)$ is retrieved by using Eq. (22).

$$f(\tau) = \sum_{\nu=0}^{\infty} F(s) e^{i2\pi(\nu/N)\tau} \tag{22}$$

See INFORMATION THEORY; Z TRANSFORM.

Hartley transform. Where a data set consists of real values and there is no point in going to complex representation, the Hartley transform offers an alternative to the Fourier transform, especially in its

TABLE 2. Theorems for the Fourier transform

Theorem	Function, $f(x)$	Fourier transform, $F(s)$
Similarity	$f(ax)$	$\lvert a\rvert^{-1}F(s/a)$
Addition	$f(x) + g(x)$	$F(s) + G(s)$
Shift	$f(x - a)$	$e^{-i2\pi as}F(s)$
Convolution	$f(x) * g(x)$	$F(s)G(s)$
Modulation	$f(x) \cos 2\pi f_0(x)$	${}^1\!/_2 F(s + f_0) + {}^1\!/_2 F(s - f_0)$
Autocorrelation	$\int_{-\infty}^{\infty} f^*(u)f(u + x)du$	$\lvert F(s)\rvert^2$
Derivative	$f'(x)$	$i2\pi sF(s)$
Rayleigh's	$\int_{-\infty}^{\infty} \lvert f(x)\rvert^2 = \int_{-\infty}^{\infty} \lvert F(s)\rvert^2 ds$	

discrete form. The definition is Eq. (23), and the inversion formula is Eq. (24), where cas $\theta = \cos\theta +$

$$H(\nu) = N^{-1} \sum_{\tau=0}^{N-1} f(\tau)\,\mathrm{cas}\,(2\pi\nu\tau/N) \qquad (23)$$

$$f(\tau) = \sum_{\nu=0}^{N-1} H(\nu)\,\mathrm{cas}\,(2\pi\nu\tau/N) \qquad (24)$$

$\sin\theta$. For computing purposes it is convenient not to handle imaginaries, especially with multidimensional applications such as images, and there is a simplicity about calling a Hartley transform subprogram which, unlike the Fourier, is the same whether the operation performed is going to the transform domain or returning from it. The power spectrum, which is computed from the Fourier transform by summing the squares of the real and imaginary parts, is obtained as $[H(\nu)]^2 + [H(N-\nu)]^2$. The complex Fourier transform of real data contains a twofold redundancy, since $F(N-\nu)$ is not independent of $F(\nu)$, and this redundancy imposes a time penalty on computation. Nonreciprocal transforms accepting complex input and delivering real output and others converting real to complex can be combined into packages to avoid the waste time, but the elegant Hartley technique has made such devices unnecessary.

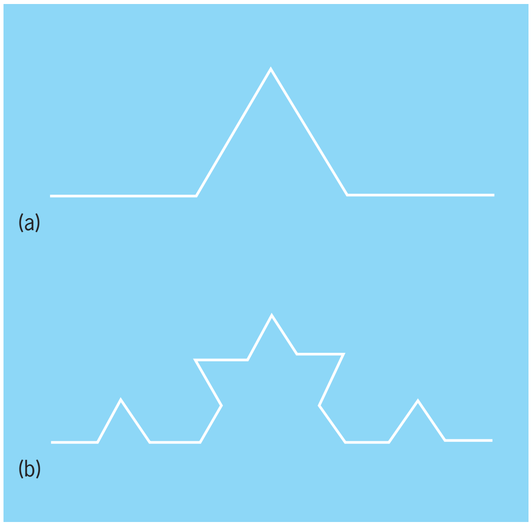
Fast Fourier transform. In applications of harmonic analysis the discrete Fourier transform may be computed directly from Eq. (21) but, where speed is essential, less time will be required if the data values are split into two sets of length $N/2$ and two separate (but shorter) discrete Fourier transforms are computed and suitably combined. A further time saving occurs if the shorter data sets are themselves halved, and so on. Because of this, data sets are often set up so that, $N = 2^P$, where P is some power, so that halving can continue down to length 2. Around 1805 Gauss was already practicing this technique and stated that an advantage could be obtained, wherever there were factors such that $N = n_1 n_2$, by dividing the data sets into lengths n_1 and n_2. An algorithm that uses this idea is called a fast Fourier transform. Significant further time saving results when $N = 4^P$, in which case the algorithm is called a radix-4 fast Fourier transform. When the number of data values given falls short, for example, if $N = 730$, it may be advantageous to append 294 zeros to the data in order to bring N up to $4^5 = 1024$. It is true that in principle the factorization $2 \times 5 \times 73 = 730$ could lead to time saving, but it is inconvenient to maintain special computer code for prime factors such as 73. Furthermore, extending with zeros means that the discrete Fourier transform values, when computed, will be packed a little closer together than is strictly necessary, which leads to a smoother graph, in the event that a power spectrum, such as an infrared spectrum, is to be displayed. In addition, the speed of computers has repeatedly doubled, whereas improvements in fast algorithms have slackened. Consequently, radix-4 algorithms can be recommended for general-purpose harmonic analysis. When data values are real, as is usually the case, the fast Hartley transform algorithm, which is strictly analogous to the fast Fourier transform, will be twice as fast as the comparable complex fast-Fourier transform computer code. Ronald N. Bracewell

Bibliography. R. N. Bracewell, *The Fourier Transform and Its Applications*, 3d ed., 1999; R. N. Bracewell, *The Hartley Transform*, 1986; R. N. Bracewell, *Two-Dimensional Imaging*, 1994; J. W. Goodman, *Introduction to Fourier Optics*, 2d ed., 1996; J. S. Walker, *Fourier Analysis*, 1988.

Fractals

Geometrical objects that are self-similar under a change of scale, for example, magnification. The concept is helpful in many disciplines to allow order to be perceived in apparent disorder. For instance, in the case of a river and its tributaries, every tributary has its own tributaries so that it has the same structural organization as the entire river except that it covers a smaller area. The branching of trees and their roots as well as that of blood vessels, nerves, and bronchioles in the human body follows the same pattern. Other examples include a landscape with peaks and valleys of all sizes, a coastline with its multitude of inlets and peninsulas, clouds, lightning, the mass distribution within a galaxy, the distribution of galaxies in the universe, and the structure of vortices in a turbulent flow. The rise and fall of economic indices has a self-similar structure when plotted as a function of time.

Mathematical versus natural fractals. A triadic Koch curve (see **illus.**) is a good example of how a fractal may be constructed. The procedure begins with a straight segment. This segment is divided into three equal parts, and the (single) central piece is replaced by two similar pieces (illus. *a*). The same procedure is now applied to each of the four new segments (illus. *b*), and this is repeated an infinite number of times. The curve is self-similar, because a magnification by 3 of any portion will look the same as



Koch curve, (*a*) first and (*b*) second stages.

the original curve. As a geometrical curve, the Koch curve has many unsettling properties. For instance, it has an infinite number of zigzags between any two points, and its length between any two points is infinite. The curve was constructed in the late nineteenth century to demonstrate that continuous curves are not necessarily differentiable and rectifiable. The invention of many such curves around that time marks the origin of fractals as a branch of mathematics.

Fractals came into natural sciences when it was recognized that natural objects are often random versions of mathematical fractals. They are self-similar in a statistical sense; that is, given a sufficiently large number of samples, a suitable magnification of a part of one sample can be matched closely with some member of the ensemble. Unlike the Koch curve, which must be magnified by an integral power of 3 to achieve self-similarity, natural fractal objects are usually self-similar under arbitrary magnification. Also, the scale factor need not be the same in every direction. The landscape is self-similar when it is magnified by one factor in the horizontal and another factor in the vertical direction. Such objects, including rough solid surfaces, are said to be self-affine fractals.

Natural fractals are self-similar or self-affine within a finite range of length scales. A coastline is limited from above by the size of the continent, and it cannot be subdivided beyond a grain of sand. Fractal properties of solid surfaces cannot be extended beyond the size of an atom. It is believed that natural fractals, like the Koch curve, are created by stacking up some basic building blocks—sand grains or atoms—following a set of simple rules under random conditions. There are many examples to illustrate this principle, but finding the rules of construction for a given fractal object is not always straightforward.

The multitude of zigzags of the Koch curve mimics the irregularities of a coastline with protrusions and indentations of sizes ranging from hundreds of kilometers down to a grain of sand. The length of such a curve depends on the scale of measurement, which in practice is the limit of resolution. For example, if the scale of the Koch curve in the illustration is the length of the original line segment, the central part of the original segment (illus. *a*) will not be visible. An improvement of the resolution by a factor of 3 makes visible this entire structure (illus. *a*) but not the finer details (illus. *b*). The curve now appears longer by a factor of $^4/_{\sqrt{3}}$. By repeating this procedure, the length increases by a factor of $^4/_{\sqrt{3}}$ every time the resolution is improved by a factor of 3. Similarly, the length of a coastline measured from a satellite picture is considerably shorter than that measured from the ground.

Fractal dimension. A magnification of a line segment by a factor of 3 results in an increase in length by the same factor. For a square, a magnification of each side by a factor of 3 results in an increase in area by a factor of 9 or 3^2. For a cube, the increase in volume is proportional to the third power of the magnification factor of each side. In every case the dimension of the object appears as the power of the length scale factor. In contrast, the length of the Koch curve as defined in the preceding paragraph increases by a factor of 4 when it is magnified by a factor of 3. An analogous definition of the dimension, $4 = 3^D$, leads to the fractal or Hausdorff dimension of the curve, $D = (\ln 4)/(\ln 3) \simeq 1.26$. The fractal dimensions of a typical coastline is 1.2. Self-affine fractals have well-defined dimensions when one length scale dominates. Landscapes and catalytically active substrates have $D \simeq 2.1$ compared with $D = 2$ for smooth surfaces. Thus, D is a quantitative measure of how much more detail is visible in an object when it is magnified.

Some fractal objects appear less substantial when examined in more detail. For instance, the distribution of matter in a galaxy is such that more empty space is apparent under higher resolution. The dimensions of such objects are less than the spatial dimension in which they are embedded, for example, 1.2 for typical galaxies. Polymer molecules in solutions form loosely packed spheres with $D \simeq 1.7$. On the other hand, space-covering fractals have dimensions equal to the spatial dimension. The river system has $D = 2$, meaning that a small body of water drains the entire land. The blood vessel system, with $D = 3$, allows a small amount of blood to carry nutrients to every cell of the body.

Physical properties. In a diffusion process in uniform space, the mean-square-average distance covered by the diffusing substance is proportional to the lapsed time. The same process on a fractal network, for example, a percolation cluster, gives a very different result. A percolation network is a lattice constructed of a random mixture of conducting and nonconducting links. At low concentrations of conducting links, the lattice is an insulator because there is no connected path between two opposite ends of the lattice. At a critical concentration, a fractal connected path is formed and the cluster becomes a conductor. In a diffusion process on a two-dimensional model of the percolation cluster, the mean-square-average distance is proportional to the time raised to the power 0.7. This is one illustration of the unusual dynamical properties of fractal objects. A fractal cluster consisting of masses connected by springs has an unusual vibrational spectrum in that the distribution of normal modes is fractal. The relations between fractal geometry, turbulent flow, and chaotic motion of nonlinear systems are under active investigation. *See* CHAOS; DIFFUSION.

General applications. Physicists have used the concept of fractals to study the properties of amorphous solids and rough interfaces and the dynamics of turbulence. It has also been found useful in physiology to analyze the heart rhythm and to model blood circulation, and in ecology to understand population dynamics. In computer graphics it has been shown that the vast amount of information contained in a natural scene can be compressed very effectively by identifying the basic set of fractals therein together with their rules of construction. When the fractals are reconstructed, a close approximation of the original scene is reproduced. *See* AMORPHOUS SOLID; CARDIOVASCULAR SYSTEM; COMPUTER GRAPHICS; TURBULENT FLOW.

S. H. Liu

Applications in physics. As noted above, natural fractal objects show structures at all scales and are therefore intrinsically irregular. Their properties cannot be described by the usual mathematical concepts that focus instead on regularity and analyticity. Fractal geometry has provided a mathematical tool to characterize these structures in a quantitative way.

Before the application of fractals to physics, scale invariance had already appeared in the study of the critical properties of continuous phase transitions. This concept was essential in the development of the renormalization group theory that has provided an understanding of the origin of scale invariance in these systems. However, in critical phenomena fractal properties appear only at the critical point and, in order to be observed, they require an extremely accurate fine tuning of an external parameter such as the temperature. *See* CRITICAL PHENOMENA; RENORMALIZATION.

By contrast, most of the fractal structures observed in nature arise from irreversible dynamical processes and do not require any fine tuning. These differences are of great conceptual importance and imply that new theoretical concepts are necessary for the understanding of these phenomena. In particular, it is essential to focus on the temporal evolution and the dynamical aspects of these phenomena, while in the study of phase transitions only the static configurations are considered.

Fractal growth models. The first model of irreversible fractal growth based on a well-defined physical process is the diffusion-limited aggregation model introduced in 1981 as a model of smoke aggregation and colloid formation. The process starts with a seed particle, and a second particle is introduced into the system which moves with random-walk dynamics (brownian motion) until it touches the first particle. At this point it stops and becomes part of the structure. A new particle is then introduced, and the process is repeated iteratively. The result of this simple process is a random fractal structure of great complexity. *See* BROWNIAN MOVEMENT.

A generalization of diffusion-limited aggregation, the dielectric breakdown model, shows that a system which obeys the Laplace equation and follows a stochastic iteration scheme spontaneously configures its boundaries into a highly irregular fractal structure. Such growth models can describe the essential fractal properties of a great variety of phenomena such as electrochemical deposition, dielectric breakdown, viscous fingering, the propagation of fractures, colloids, and polydisperse solutions. *See* LAPLACE'S DIFFERENTIAL EQUATION; STOCHASTIC PROCESS.

Unsolved problems. Nevertheless, in many other systems well-defined fractal properties have been identified either from experiments or by computer simulations, but theoretical understanding is lacking. An important example is fully developed turbulence, for which there is empirical and phenomenological evidence that energy dissipation occurs on a highly irregular, possibly fractal structure, but the connection of this property to the Navier-Stokes equations is still unknown.

Another important problem is represented by the large-scale structure of the universe. Visible matter is aggregated in galaxies that are part of groups that are aggregated into clusters and superclusters, and so on. The analysis of galaxy and cluster catalogs with the methods of modern statistical mechanics shows that visible matter has fractal properties up to the present limits of observation. This situation is in contrast with the usual assumption of large-scale homogeneity and leads to a theoretical challenge. *See* COSMOLOGY; GALAXY, EXTERNAL; UNIVERSE.

Disordered media. A more natural and direct application of fractal concepts is in the field of disordered systems and porous media, mentioned above. Often these systems show well-defined fractal behavior, so it is important to study the physical properties of fractal structures such as their vibrations (fractons) and their electronic and transport properties.

L. Pietronero

Bibliography. M. Barnsley, *Fractals Everywhere*, 2d ed., 1993; P. H. Coleman and L. Pietronero, The fractal structure of the universe, *Phys. Rep.*, 213: 311–391, 1992; H. Ehrenreich and D. Turnbull (eds.), *Solid State Physics*, vol. 39, 1986; B. B. Mandelbrot, *The Fractal Geometry of Nature*, 1988; L. Pietronero and E. Tosatti (eds.), *Fractals in Physics*, 1986; T. Vicsek, *Fractal Growth Phenomena*, 1992.

Frame of reference

A base to which to refer physical events. A physical event occurs at a point in space and at an instant of time. Each reference frame must have an observer to record events, as well as a coordinate system for the purpose of assigning locations to each event. The latter is usually a three-dimensional space coordinate system and a set of standardized clocks to give the local time of each event. For a discussion of the geometrical properties of space-time coordinate systems *see* SPACE-TIME; RELATIVITY.

Newtonian reference frames. In the ordinary range of experience, where light signals, for all practical purposes, propagate instantaneously, the time of an event is quite distinct from its space coordinates, since a single clock suffices for all observers, regardless of their state of relative motion. The set of reference frames which have a common clock or time is called newtonian, since Isaac Newton regarded time as having invariable significance for all observers.

Inertial reference frames. In these a body moves with constant velocity when free of impressed forces. It follows that one inertial frame cannot be accelerating with respect to another. If x_1, y_1, z_1 and x_2, y_2, z_2 are the coordinates of a particle P in parallel inertial frames 1 and 2, where 2 is moving with respect to 1 with velocity components v_x, v_y, v_z, then Eqs. (1) hold. (Newtonian frames are understood

$$\begin{aligned} x_2 &= x_1 - v_x t \\ y_2 &= y_1 - v_y t \\ z_2 &= z_1 - v_z t \end{aligned} \tag{1}$$

here so that $t_2 = t_1$.) This is called a galilean transformation. Since the vs are independent of time, the component accelerations of P in 1 and 2 are the same, that is, Eq. (2) is valid.

$$A_{x_2} = \frac{d^2x_2}{dt^2} = \frac{d^2x_1}{dt^2} = A_{x_1} \text{ etc.} \qquad (2)$$

Accelerated reference frames. A noninertial frame of reference is called an accelerated frame. A point fixed in such a frame will be accelerated with respect to inertial frames and so must be held by a force. An observer in an accelerated frame will see accelerations that are not proportional to the impressed forces, since a force is required to keep the particle at rest. Newton's second law does not hold without modification. *See* ACCELERATION.

Rotating reference frames. A rotating frame of reference is one whose coordinate axes are rotating with respect to some inertial system. A point P fixed in the rotating frame moves uniformly about a circle in the inertial frame and is accelerated toward the axis of rotation. There is an additional acceleration (with respect to the inertial system) if P moves in the rotating system. For example, if the point moves away from the axis along a moving radius, it gains circular speed also. *See* CENTRIPETAL FORCE; CORIOLIS ACCELERATION.

Newton's second law can be modified to take account of these two accelerations that are not observed in the rotating system. Let A_{Ac}, A_C, A_{Co} be the observed, centripetal, and Coriolis accelerations, respectively. The total acceleration in the inertial frame is given by Eq. (3), where the accelerations must

$$A = A_{Ac} + A_C + A_{Co} \qquad (3)$$

be added vectorially. Defining the centrifugal force, $Fc = -mAc$, and the Coriolis force, $F_{Co} = -mA_{Co}$, permits Newton's second law to be written as Eq. (4).

$$mA_{Ac} = F + F_C + F_{Co} \qquad (4)$$

In rotating systems a consistent mechanical theory is obtained if the two additional forces are added to the impressed forces. Conversely, the presence of such forces will enable the observer to determine the motion of his reference frame.

Astronomical reference frames. These are fixed relative to celestial objects. The most convenient one for observational purposes is, of course, a geocentric system. The Earth's axis of rotation is used as a coordinate axis. A second reference direction is chosen as the intersection of the Earth's equatorial plane and the plane of the Earth's orbit (the ecliptic). The basic time unit is the Earth's rotation period, the sidereal day. The positions and times of distant astronomical events are determined indirectly. *See* ASTRONOMICAL COORDINATE SYSTEMS.

Reference frames are regarded as more basic as more phenomena appear regular with respect to them. Thus the planetary motions are described naturally in an ecliptic coordinate system centered on the Sun, and the motions of stars as viewed from the Earth appear regular in a coordinate system with its origin at the center of the galaxy.

Reference frames based on the mean positions of large enough celestial objects, for example, a group of galaxies, approximate inertial frames closely. Between two such reference frames, the distances and relative velocities may be so large that transmission of light signals takes considerable time. Newtonian reference frames with a single universal time are no longer meaningful, and each reference frame must have its own system of time. Inertial frames of this type are called galilean reference frames, a somewhat misleading name since the transformation of the coordinates of an event from one such frame to another is not the Galilean transformation given earlier (with $t_1 = t_2$) but one which involves the time directly. It is known as a Lorentz transformation. Galilean reference frames are often called Lorentz frames. *See* LORENTZ TRANSFORMATIONS. Bernard Goodman

Bibliography. A. Einstein, *The Meaning of Relativity*, 5th ed., 1956; D. Halliday, R. Resnick, and K. S. Krane, *Physics*, 2 vols., 5th ed., 2002; J. B. Hartle, *Gravity*, 2003; C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, Berkeley Physics Course, vol. 1, 2d ed., 1973; D. Morrison, S. C. Wolff, and A. Fraknoi, *Abell's Exporation of the Universe*, 7th ed., 1995.

Francisella

A genus of very small, coccoid to ellipsoidal, pleomorphic, nonmotile and nonsporulating, gram-negative, rod-shaped bacteria. It was named for Edward Francis, who studied the tularemia bacterium. Fastidious and strictly aerobic, it grows at 98.6°F (37°C) within several days, but only in enriched media such as coagulated egg yolk or glucose-cysteine-blood-agar. The genus is catalase-negative. Acid is produced from several carbohydrates; indole, urease, gelatinase, and nitrate reductase are not formed. The genome deoxyribonucleic acid contains 32–36 mol % guanine plus cytosine. The organisms occur in natural waters of the Northern Hemisphere, and can be parasitic and pathogenic in birds, anthropods, and mammals, including humans.

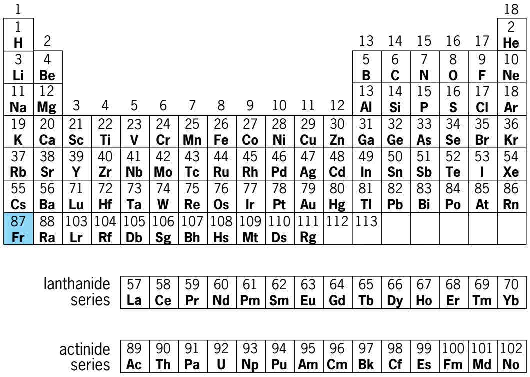
Francisella tularensis was first isolated in 1911 from ground squirrels with a plaguelike disease in Tulare County, California. It has since been found in many wild animals throughout North America, continental Europe, and Asia. Tularemia in humans is acquired via transmission by blood-sucking arthropods or by contact with infected animals, most frequently by hunters and butchers. *See* TULAREMIA.

Francisella novicida was isolated only once from a water pond, and various other *Francisella*-like organisms have been isolated from patients but so far have remained unclassified. *See* ANTIBIOTIC; MEDICAL BACTERIOLOGY. Walter Mannheim

Bibliography. N. R. Krieg and J. G. Holt (eds.), *Bergey's Manual of Systematic Bacteriology*, 9th ed., vol. 1, 1984; M. P. Starr (ed.), *Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria*, 1989.

Francium

A chemical element, Fr, atomic number 87, an alkali metal element falling below cesium in group 1 of the periodic table. Distinguished by nuclear instability, francium exists only in short-lived radioactive forms, the most durable of which has a half-life of 21 min. The chief isotope of francium is actinium-K, an isotope of mass 223, which arises from the radioactive decay of the element actinium. From the properties of the known isotopes, it is reasonably certain that no long-lived form of element 87 will ever be found in nature or synthesized artificially. *See* PERIODIC TABLE.



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The chemical properties of francium can be studied only on the tracer scale. The element has all the properties expected of the heaviest alkali element. With few exceptions, all the salts of francium are water-soluble. *See* ACTINIUM; ALKALI METALS.

Earl K. Hyde

Bibliography. G. Choppin et al., *Radiochemistry and Nuclear Chemistry*, 2001; G. Friedlander et al., *Nuclear and Radiochemistry*, 3d ed., 1981; E. K. Hyde, *Radiochemistry of Francium*, U.S. Atomic Energy Commission, 1960.

Franck-Condon principle

The generalization that the transition from one energy state of a molecular system to another occurs so nearly instantaneously that the nuclei of the atoms involved can be considered as stationary during the process. The Franck-Condon principle is closely related to the Born-Oppenheimer approximation, in which the various motions (electronic, nuclear vibrations and rotations) are considered to be separable, and in which the electrons respond to the instantaneous vibrations of the system whereas the system responds only to the average position of the electrons. The principle, proposed by J. Franck in 1925 and developed quantum-mechanically by E. U. Condon in 1928, is important in discussing systems of more than one atom. It is therefore valuable in molecular spectroscopy and in the interpretation of the optical properties of liquids and solids.

The principle may be illustrated by considering the excitation of photoluminescence in molecules or in special luminescent sites (centers) in solids. In the state of lowest electronic energy (ground state), the interatomic distance in a diatomic molecule or the distance of a given atom in a center from its nearest neighbors (a distance called a configuration coordinate) so adjusts itself that the total energy of the system assumes a minimum value. Within this ground electronic state, however, the system may occupy many vibrational energy states, depending on the motion of the atoms composing the molecule or center. In a higher electronic state, where the forces involved are different from those in the ground state, the system has a different set of vibrational levels and a different minimum-energy interatomic distance or configuration coordinate.

If the luminescent system is in its ground state and absorbs light which raises it to an electronically excited state, the Franck-Condon principle asserts that immediately after the transition the nuclei are still in the ground-state equilibrium configuration. This configuration does not represent the condition of minimum energy for the excited state, but usually corresponds to this state in a condition of high vibrational excitation. Since the vibrational frequency of molecules or centers is about 10^{13}/S whereas the transition probability for an allowed transition emitting visible light is about 10^{8}/S, the molecule or center has ample time to relax to the new minimum configuration characteristic of the electronically excited state. It does this by giving off the excess vibrational energy as heat.

A similar process occurs in the luminescence transition back to the electronic ground state, where heat energy is again given off when the luminescent center readjusts to the ground-state equilibrium configuration and the center returns to its initial state. However, the absorption transition is achieved by the action of light energy only, whereas the system returns to the ground state by giving up heat energy twice and emitting the remaining energy as luminescence. As a result, the luminescent emission is of lower energy than the absorbed light, and since the wavelength varies inversely as the energy, the emission is of longer wavelength than is the absorbed light. This general result, known as the Stokes shift in luminescence, is thus explained on the basis of the Franck-Condon principle. *See* LUMINESCENCE; MOLECULAR STRUCTURE AND SPECTRA.

James H. Schulman; Clifford C. Klick

Bibliography. F. Argullo-Lopez and C. R. Catlow, *Point Defects in Materials*, 1988; J. H. Crawford and L. M. Slifkin (eds.), *Point Defects in Solids*, vol. 1, 1972; R. J. Elliott and A. F. Gibson, *An Introduction to Solid State Physics*, 1974.

Fraunhofer lines

Dark absorption features in the solar spectrum. They are named in honor of J. Fraunhofer, who first studied them in 1814. They are found from the ultraviolet at about 180 nanometers to the infrared at 20 micrometers. Each line represents the net

Fraunhofer lines

Wavelength, nm	Name*	Species†	Cycle variability,‡ %	Comment§
279.54		Mg II	10	Ultraviolet emission, high chromosphere
280.23				
388.36	(CN bandhead)	CN	3	High photosphere
393.36	K	Ca II	15	Chromosphere
396.85	H			
486.13	F (Hβ)	H I		Chromosphere
517.27	b_2	Mg I	—	Low chromosphere
518.36	b_1			
525.02		Fe I	0.3	Photosphere, magnetic fields ($g = 3$)
538.03		C I	0.0	Low photosphere
589.00	D_2	Na I	—	Upper photosphere, low chromosphere
589.59	D_1			
630.25		Fe I	—	Photosphere, magnetic fields ($g = 2.5$)
656.28	C (Hα)	H I	6	Chromosphere
849.80		Ca II	1	Low chromosphere
854.21				
866.22				
868.86		Fe I	—	Photosphere, magnetic fields ($g = 1.7$)
1083.03		He I	200	High chromosphere
1281.8	Paschen β	H I	—	Chromosphere
1564.85		Fe I	—	Photosphere, magnetic fields ($g = 3$)
4665		CO	—	High photosphere
12,320		Mg I	—	High photosphere, magnetic fields ($g = 1$)

[*] J. Fraunhofer's original designation, if any.
[†] Atom or molecule responsible.
[‡] Any temporal variability observed in integrated light.
[§] Origin and diagnostic value of the line for solar physics research.

absorption of light by a particular atom or molecule. Most form in the Sun's atmosphere, although the Earth's telluric spectrum contributes lines of molecular oxygen (O_2), carbon monoxide (CO), and other molecules. Some lines such as Fraunhofer's C line in the red (hydrogen-alpha) can be seen with a pocket spectroscope. Powerful research instruments reveal millions of lines, most of which are weak and blended together in an almost inextricable tangle. The **table** lists some interesting Fraunhofer lines.

A spectrum line is caused by the absorption of photons of light that excite the atom from a lower to a higher energy level. Spontaneous decay back to the atom's lower level then follows, accompanied by the isotropic emission of light at the wavelength of the line. The result is a loss of light in the Sun-Earth direction. The probability of a given photon being absorbed depends on the quantum state of the atom.

The study of the Fraunhofer spectrum is the principal means of learning about physical conditions in the solar atmosphere. On the resolved solar disk, variations in line strength form point to point convey information about temperature, Doppler shifts of the lines reveal gas motions, and line splitting from the Zeeman effect maps magnetic fields. Because each line represents a chemical element, the composition of the solar atmosphere can be deduced. *See* DOPPLER EFFECT; SOLAR MAGNETIC FIELD; SUPERGRANULATION; ZEEMAN EFFECT.

Profiles. A high-dispersion spectrograph shows that Fraunhofer lines are not sharp; most have a gaussian profile with extended wings. This broadening results mainly from the Doppler motions associated with thermal and unresolved turbulence, together with collisional damping, wherein the atoms interact with their neighbors during the excitation-deexcitation process. Like line intensity and wavelength shifts, line profiles can yield further information about the solar surface, for example, the gas density and pressure. *See* SPECTROGRAPH.

Variability. The strength of the weak line of carbon at 538.0 nm is a sensitive indicator of surface temperature. The cores of chromospheric calcium H and K lines at 396.8 nm and 393.3 nm, which originate in active regions, correlate with the Sun's ultraviolet output. A variability of the Sun that might affect the Earth can be looked for by measuring the temporal variation of these lines in integrated sunlight. *See* ASTRONOMICAL SPECTROSCOPY; SOLAR CONSTANT; SUN.

William C. Livingston

Bibliography. A. Bhatnagar and W. C. Livingston, *Fundamentals of Solar Astronomy*, 2005; P. V. Foukal, *Solar Astrophysics*, 2d ed., 2004; L. Golub and J. M. Pasachoff, *Nearest Star: The Surprising Science of Our Sun*, 2001; K. H. Schatten and A. Arking (eds.), *Climate Impact of Solar Variability*, 1990.

Free-electron theory of metals

The treatment of a metal as containing a gas of electrons completely free to move within it. The theory was originally proposed in 1900 to describe and correlate the electrical and thermal properties of metals. Later, quantum mechanics became the basis for the theory of most of the general properties of simple metals such as sodium, with one free electron per atom, magnesium with two, and aluminum with three. Transition metals, such as iron, have partially filled electronic *d* states and are not treated by the free-electron model.

Development. Three years after J. J. Thomson's 1897 discovery of the electron, P. Drude suggested that the transport properties of metals might be understood by assuming that their electrons are free and in thermal equilibrium with their atoms. This theory was made more quantitative by H. A. Lorentz. Assuming that the mean free path of electrons was limited by collisions, he was able to derive Ohm's law for the electrical conductivity and obtain the ratio of thermal to electrical conductivity in excellent agreement with experiment. This ratio, divided by the absolute temperature, is called the Wiedemann-Franz ratio and had been observed to be universal 50 years earlier. *See* CONDUCTION (ELECTRICITY); CONDUCTION (HEAT); ELECTRICAL CONDUCTIVITY OF METALS; KINETIC THEORY OF MATTER; MEAN FREE PATH; OHM'S LAW; THERMAL CONDUCTION IN SOLIDS; WIEDEMANN-FRANZ LAW.

The theory, however, had two major shortcomings. First, it predicted a large component of the specific heat of a metal, not present in insulators, which was not observed. Second, comparison of the theory with experiment indicated that the mean free path of the electrons became extremely large at low temperatures; the model offered no justification.

In 1928 A. Sommerfeld revised Lorentz's treatment by using quantum statistics, which removed the difficulty of the specific heat without losing the successful description of transport properties. The resulting theory remains the basis for the understanding of most transport properties of metals and semiconductors. At about the same time, W. V. Houston and F. Bloch solved the quantum-mechanical wave equation for electrons in a regular periodic structure, finding that they could indeed have arbitrarily large mean free paths if there were no defects in the periodicity, thereby putting the free-electron theory on a firm basis. *See* BAND THEORY OF SOLIDS; BLOCH THEOREM; FERMI-DIRAC STATISTICS; SPECIFIC HEAT OF SOLIDS; STATISTICAL MECHANICS.

Equilibrium properties. In the classical theory of Drude and Lorentz, the state of an electron can be specified by giving its position **r** and momentum **p**. The corresponding state-space, the six-dimensional **r, p** space, is called phase space. For the statistical analysis of the electron gas, it was convenient to divide the phase space into cells and specify the number of electrons in each cell as a function of time. In quantum theory the electrons are also treated as waves, satisfying appropriate conditions at the boundaries of the metal. This limits the wavelengths λ, and therefore the momentum values $p = h/\lambda$, where h is Planck's constant. It is easily shown that there is one allowed state for each cell of volume h^3 in phase space. The Pauli exclusion principle allows each state to contain only a single electron with each of the two spins. Thus in quantum theory the number of electrons in each cell is replaced by the probability $f(\mathbf{p}, \mathbf{r}, t)$ that the cell at the momentum **p** and position **r** is occupied by an electron of spin-up, usually equal to the probability of occupation by an electron of spin-down. The function $f(\mathbf{p}, \mathbf{r}t)$ is called the distribution function for the electrons. *See* EXCLUSION PRINCIPLE.

For a free-electron gas in thermal equilibrium, the distribution function does not depend upon time and is the same at all **r**. Furthermore, it does not depend upon the direction of motion and is thus a function only of the magnitude of the momentum p, or the energy $E = p^2/2m$, where m is the electron mass. For a given number of electrons and a given total energy, there are many ways in which the electrons can be distributed in phase space. Statistical mechanics provides the distribution function $f_0(E)$, for a given number of electrons and given total energy, that can be obtained in the most ways. The subscript zero indicates the equilibrium distribution, which is the Fermi-Dirac distribution function given by Eq. (1). Here E_F is called the Fermi energy

$$f_0(E) = \frac{1}{e^{(E-E_F)/k_BT} + 1} \qquad (1)$$

and for any system takes a value giving the correct total number of electrons; k_B is the Boltzmann constant; and the temperature T determines the total energy.

If the number of free electrons is very small, E_F is far below the energy of the lowest state, $E = p^2/2m = 0$, and the exponential is so large that the number one in the denominator can be neglected, and the classical Boltzmann distribution proportional to e^{-E/k_BT} is obtained, corresponding to Drude's theory. However, in real metals the density of electrons is so high that states must be occupied up to energies of 5–10 eV, whereas k_BT at room temperature is 0.025 eV, and the opposite limit is appropriate. Then E_F becomes of the order of 5 or 10 eV; $f_0(E)$ is near one for E up to that energy and then drops, over a range of a few hundredths of an electronvolt, to approximately zero. Consistent with the Pauli exclusion principle, almost all levels are occupied up to about the Fermi energy, and almost none above. Changing the temperature has only a small effect on the total energy; the large classical contribution, $^3/_2\, k_B$ per electron, to the specific heat is eliminated. In fact, only those electrons near the Fermi energy, a fraction of order k_BT/E_F, can take up thermal energy; the specific heat is $\pi^2 k^2_B T/2E_F$ per electron (a $\pi^2/3$ factor was derived from the full analysis), consistent with experiment. *See* BOLTZMANN STATISTICS.

An understanding of magnetic susceptibility in free-electron theory requires quantum mechanics. The equilibrium distribution for electrons in the presence of a magnetic field is needed. However, since the field deflects electrons in classical physics without energy change, the equilibrium distribution is unaffected; the energy is independent of field and the susceptibility is zero. In quantum theory the deflected electrons form orbits, which are quantized as are orbits in the atom. When these quantized levels are statistically occupied, the total energy does depend upon the field and a diamagnetic susceptibility is predicted. This contribution, along with a contribution from the partial alignment of electron spins, describes the observed susceptibility of metals. *See* DIAMAGNETISM; MAGNETIC SUSCEPTIBILITY.

Transport theory. When electric fields are applied, the electrons are no longer in equilibrium. For a steady voltage applied to a long wire, the distribution function is still independent of time and position along the wire, but depends upon the direction of the electron momentum, $f(\mathbf{p})$. In free-electron theory the distribution function can be determined by following $f[\mathbf{p}(t)]$ as the momentum changes under the influence of the electric field $\mathscr{E}$ according to Eq. (2), where $-e$ is the electronic charge. The probability of the occupation of the state would not change with time except for electron collisions which can scatter the electron to another momentum, or scatter another electron to this momentum; thus Eq. (3) holds. The distribution function is assumed to be close to the equilibrium distribution written in Eq. (4), where $f_1(\mathbf{p})$ is small and proportional to the electric field $\mathscr{E}$.

$$\frac{d\mathbf{p}}{dt} = -e\mathscr{E} \tag{2}$$

$$\frac{df[\mathbf{p}(t)]}{dt} = \left.\frac{\partial f(\mathbf{p})}{\partial t}\right|_{\text{collisions}} \tag{3}$$

$$f(\mathbf{p}) = \frac{f_0 p^2}{2m} + f_1(\mathbf{p}) \tag{4}$$

Thus evaluation of the left side of Eq. (3) in terms of patial derivatives produces a term $df_0(E)/dE$ $(dF/d\mathbf{p}) \cdot (-e\mathscr{E})$ and a term proportional to f_1 and to $\mathscr{E}$, and therefore negligible. The collisions are expected to relax any deviation $f_1(\mathbf{p})$ from the equilibrium distribution back toward zero, more rapidly with larger deviation. Then the right side of Eq. (3) is approximated by $-f_1(\mathbf{p})/\tau$. This is called the relaxation-time approximation, and τ is the relaxation time, usually assumed to be independent of $\mathbf{p}$, $\mathbf{r}$, and t. Then Eq. (5) may be used to solve for $f_1(\mathbf{p})$, noting that $dE/d\mathbf{p}$ is the electron velocity $\mathbf{v}$. *See* RELAXATION TIME OF ELECTRONS.

$$f_1(p) = e\mathbf{v} \cdot \mathscr{E}\,\tau \frac{df_0(E)}{dE} \tag{5}$$

This analysis is the derivation and solution of the linearized Boltzmann equation in the relaxation-time approximation. The distribution having been found, the current contribution $-e\mathbf{v}$ can be summed for the state in each cell, multiplied by the probability that it is occupied, and by two for spin, and divided by the volume to obtain the current density. The sum over cells is written in Eq. (6) as an integral over phase space, $(1/h^3)\Omega\int d^3p$, where Ω is the volume. The last step requires a partial integration to obtain a form like that for the electron density N. The proportionality factor $Ne^2\tau/m$ is the predicted conductivity σ and is the same as that obtained for the Boltzmann distribution assumed by Drude and Lorentz.

$$\mathbf{j} = \frac{2}{h^3\Omega}\Omega \int d^3p\,(-e\mathbf{v})f_1(p) = \frac{Ne^2\tau}{m}\mathscr{E} \tag{6}$$

In a similar way a temperature gradient is accommodated by allowing the T in Eq. (1) to vary along the wire and summing the flow of energy $\mathbf{v}p^2/2m$ for each occupied state to obtain the heat current. With an electric field strong enough to prevent any electric current, the heat current yields the thermal conductivity, as in Eq. (7). The Wiedemann-Franz ratio is a constant, $\kappa/(T\sigma) = \pi^2 k^2_B/(3e^2)$, as predicted by the classical theory, but with a different value; the factor $\pi^2/3$ arises because only electrons near the Fermi energy contribute.

$$\kappa = \frac{\pi^2}{3}\frac{k_B^2}{e^2}T\sigma \tag{7}$$

The field that arises from the temperature gradient is a thermoelectric effect. It gives rise to the voltage measured from a thermocouple and is related thermodynamically to the Peltier heat, which is generated or absorbed when current passes through a thermocouple. These equilibrium and transport properties constitute the essence of the free-electron theory of metals. The most drastic approximation may be the relaxation-time approximation, but within free-electron theory more accurate approximations are possible for the scattering term. *See* THERMOELECTRICITY.

Electron-lattice interactions. The free-electron theory neglects the interactions between the electrons and the atoms making up the metallic crystal, except for the effects of collisions in the transport theory. This was first justified by applying Floquet's theorem to the quantum-mechanical wave equation in a perfectly periodic lattice to show that the solutions can be taken in the form of perfectly periodic functions of position $u(\mathbf{r})$, modulated by free-electron waves. (The modulation is of the form of the free-electron wave function, $e^{i2\pi\mathbf{p}\cdot\mathbf{r}/h}/\sqrt{\Omega}$.) These waves indeed propagate through the perfect crystal without scattering, although the presence of a $u(\mathbf{r})$ that is not constant can modify the velocity of propagation. Such corrections could be incorporated in the free-electron theory by replacing the electron mass by an effective mass m^*, but in fact the corrections are so small in the simple metals that it is ordinarily adequate to retain the true electron mass. *See* CRYSTAL STRUCTURE.

Then it is only the deviations from periodicity, such as impurity atoms or lattice vibrations, which produce the scattering of electrons assumed in transport theory. It is then to be expected that in pure materials at low temperatures the mean free paths become extremely long. As expected, the resistivity, $\rho = 1/\sigma$, is proportional to concentration of impurities. At elevated temperatures where the lattice-vibration amplitudes are proportional to $\sqrt{k_B T}$, the scattering rate is predicted to be proportional to temperature, as observed. Thus these interactions give significance to the parameters of free-electron theory, and do not invalidate it. *See* CRYSTAL DEFECTS; ELECTRICAL RESISTIVITY; LATTICE VIBRATIONS.

It was recognized in the 1950s that the effects of the periodic lattice on the electron energies are remarkably small. The recognition was based on the observation that the momentum values $\mathbf{p}$ for the electrons that had the Fermi energy differed very little from those of the sphere in wavenumber space, $p^2/2m = E_F$, which free electrons would have. This surface in wavenumber space is called the Fermi surface. Although the potential from each atom felt by the electron is strong, its effect, called the pseudopotential, is weak. The pseudopotential is in fact so weak that it can be said only to affect electrons that satisfy the conditions for diffraction. The corresponding restructuring of the Fermi surface by diffraction leads to the nearly free-electron Fermi surfaces, which were accurately confirmed experimentally. From the point of view of free-electron theory, there may be a small number of electrons at any moment involved with diffraction, but the overwhelming majority hardly experience any potential at all and are therefore truly free. *See* FERMI SURFACE.

Interactions among electrons. Even in the context of a free-electron gas, there are strong Coulomb interactions between electrons which are frequently neglected in the free-electron theory of metals. This neglect was justified in the late 1950s by L. D. Landau, who asserted that, even with strong electron–electron interactions, there is a one-to-one correspondence between the excited states, called quasiparticle states, of the real system and the one-electron excitations from the ground state (the $T = 0$ equilibrium state) of the noninteracting electron gas. Thus, all of the formulation given here for free-electron theory still follows, but perhaps with modifications of parameters such as mass. Subsequent theory indicates that indeed these modification due to the electron–electron interaction are extremely small for the low-energy excitations present in thermal equilibrium, and so again the simplest theory succeeds for many properties, although substantial modifications are required for the higher-energy excitations caused by light. There are additional corrections, which are much larger than those from the electron–electron interaction, arising from the interaction between electrons and phonons, the quantum-mechanical term for lattice vibrations. In many metals these vibrations reduce the electron velocities by factors of as much as 2, increasing the electronic specific heat although they turn out not to modify the conductivity itself. *See* PHONON.

Another feature of the electron–phonon interaction is a resulting interaction among electrons, which is attractive and tends to cancel or exceed the repulsive electron–electron interaction. At low temperatures the net attraction binds electrons in pairs in a superconducting state. The theory of J. Bardeen, J. R. Schrieffer, and L. N. Cooper (the BCS theory of superconductivity), which first explained this phenomenon, is also a free-electron theory, but assumes that the free electrons have such a net attractive interaction. In contrast, it is generally believed that the high-temperature superconductors discovered in 1986 are very far from free-electron in character, and most workers do not believe that phonons are primarily responsible for the attractive interaction. *See* SOLID-STATE PHYSICS; SUPERCONDUCTIVITY.

Walter A. Harrison

Bibliography. J. A. Alonso and N. H. March, *Electrons in Metals and Alloys*, 1989; N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; W. A. Harrison, *Solid State Theory*, 1970, reprint 1980; C. M. Hurd, *Electrons in Metals*, 1975, reprint 1982; G. Lehman and P. Ziesche, *Electronic Properties of Metals*, 1991.

Free energy

A term in thermodynamics which in different treatments may designate either of two functions defined in terms of the internal energy U or enthalpy H, and the temperature-entropy product TS.

The function $A = U - TS$ is called the Helmholtz energy (formerly the work function or the Helmholtz free energy). The analogous function $G = H - TS$ is called the Gibbs energy (formerly the Gibbs free energy, or simply the free energy). *See* WORK FUNCTION (THERMODYNAMICS).

Theory. For a closed system (no transfer of matter across its boundaries), the work which can be done in a reversible isothermal process is given by the series shown in Eq. (1). For these conditions, $T\Delta S$

$$\begin{aligned} W_{\text{rev}} &= -\Delta A = -\Delta(U - TS) \\ &= -(\Delta U - T\Delta S) \end{aligned} \qquad (1)$$

represents the heat given up to the surroundings. Should the process be exothermal, $T\Delta S < 0$, then actual work done on the surroundings is less than the decrease in the internal energy of the system. The quantity $(\Delta U - T\Delta S)$ can then be thought of as a change in free energy, that is, as that part of the internal energy change which can be converted into work under the specified conditions. This then is the origin of the name free energy. Such an interpretation of thermodynamic quantities can be misleading, however; for the case in which $T\Delta S$ is positive, Eq. (1) shows that the decrease in "free" energy is greater than the decrease in internal energy. *See* CHEMICAL THERMODYNAMICS.

For constant temperature and pressure in a reversible process the decrease in the Gibbs energy G for the system again corresponds to a free-energy change in the above sense, since it is equal to the work which can be done by the closed system other than that associated with its change in volume ΔV under the given constant pressure P. The relations shown in Eq. (2) can be formed since $\Delta H = \Delta U + P\Delta V$.

$$\begin{aligned} \Delta G &= -(\Delta H - T\Delta S) = W_{\text{net}} \\ &= W_{\text{rev}} - P\Delta V \end{aligned} \qquad (2)$$

Each of these free-energy functions is an extensive property of the state of the thermodynamic system. For a specified change in state, both ΔA and ΔG are independent of the path by which the change is accomplished. Only changes in these functions can be measured, not values for a single state.

The thermodynamic criteria for reversibility, irreversibility, and equilibrium for processes in closed systems at constant temperature and pressure are expressed naturally in terms of the function G. For any infinitesimal process at constant temperature and pressure, $-dG \geqq \delta w_{\text{net}}$. If δw_{net} is never negative, that is, if the surroundings do no net work on the system, then the change dG must be negative or zero. For a reversible differential process, $-dG > \delta w_{\text{net}}$; for an irreversible process, $-dG > \delta w_{\text{net}}$. The Gibbs energy G thus decreases to a minimum value characteristic of the equilibrium state at the given temperature and pressure. At equilibrium, $dG = 0$ for any differential process taking place, for example, an infinitesimal change in the degree of completion of a chemical reaction. A parallel role is played by the Helmholtz energy A for conditions of constant temperature and volume. Because temperature and pressure constitute more convenient working variables than temperature and volume, the Gibbs energy is more common in thermodynamics.

Partial molal quantities. For a particular homogeneous phase in the absence of surface, gravitational, and magnetic forces, the Gibbs energy G depends on the amount of constituents present, the temperature T, and the pressure P. Let Ω represent the total number of constituents, n_i the amount of constituent i, and designate by subscript n constant composition, by subscript n_j constancy of the amounts of all constituents except n_i; then Eq. (3) is formed.

$$dG(T,P,n_1,\ldots,n_\Omega) = \left(\frac{\partial G}{\partial T}\right)_{P,n} dT + \left(\frac{\partial G}{\partial P}\right)_{T,n} dP + \sum_{i=1}^{\Omega} \left(\frac{\partial G}{\partial n_i}\right)_{T,P,n_j} dn_i \qquad (3)$$

In Eq. (3) the term $(\partial G/\partial n_j)_{T,P,nj}$ is the chemical potential μ_i of the ith constituent. It is identical to the partial molal Gibbs energy $\overline{G}_i$. It then follows that Eq. (4) holds.

$$dG = -S\,dT + V\,dP + \sum_{i=1}^{\Omega} \mu_i\,dn_i \qquad (4)$$

See SOLUTION.

Because the chemical potentials at constant T, P are intensive variables whose values are fixed, like that of the density, by relative amounts of the various constituents present, and are independent of the total mass of the phase, this equation can be integrated for constant T, P and relative composition starting from $n_i = 0$ to obtain Eq. (5). This procedure yields Eq. (6). Consistency with the expression for dG in Eq. (4) requires that Eq. (7) hold. This is the Gibbs-Duhem equation. For constant temperature and pressure, this relation imposes a condition on the composition variation of the set of chemical potentials.

$$G(T,P,n_1,\ldots,n_\Omega) = \sum_{i=1}^{\Omega} n_i\mu_i \qquad (5)$$

$$dG = \sum_{i=1}^{\Omega} \mu_i\,dn_i + \sum_{i=1}^{\Omega} n_i\,d\mu_i \qquad (6)$$

$$S\,dT - V\,dP + \sum_{i=1}^{\Omega} n_i\,d\mu_i = 0 \qquad (7)$$

Heterogeneous systems. The Gibbs energy of a closed, heterogeneous system is the sum of the Gibbs energies of its various phases. In the absence of such a constraint as provided by the subdivision of the system by a rigid, semipermeable membrane, the general thermodynamic criterion of equilibrium requires that the temperature and pressure be uniform throughout the system and that the chemical potential of each constituent have a common value for all phases in which it is present. Further, if any of the constituents can be formed from others, the chemical potentials of the reactants and products are related in accordance with the stoichiometry of the reaction equation. Thus, for the reaction in Eq. (8), at equilibrium Eq. (9) can be formed. Expressing each chemical potential μ_i in terms of the standard value μ_i° and its associated activity term $RT \ln a_i$ results in Eq. (10). In Eq. (10) $\Delta_r G^\circ$ is called the standard Gibbs energy for the reaction. Its value depends on the standard states chosen, but for a given temperature and pressure, it is a constant characteristic of the reaction involved. A true equilibrium constant K then results as shown by Eqs. (11) and (12). If the pressure for each standard state is fixed and independent of the pressure of the reaction system, $\Delta_r G^\circ$ and hence K are functions of temperature only. This is the conventional approach in treating gas-phase equilibria, but not the approach that is used ordinarily for condensed phases. *See* ACTIVITY (THERMODYNAMICS).

$$\mathrm{A} + 2\mathrm{B} \rightleftharpoons 3\mathrm{C} + 4\mathrm{D} \qquad (8)$$

$$\mu_\mathrm{A} + 2\mu_\mathrm{B} = 3\mu_\mathrm{C} + 4\mu_\mathrm{D} \qquad (9)$$

$$RT \ln \left(\frac{a_\mathrm{C}{}^3 a_\mathrm{D}{}^4}{a_\mathrm{A} a_\mathrm{B}{}^2}\right)_{\text{equil}} = -(3\mu_\mathrm{C}{}^\circ + 4\mu_\mathrm{D}{}^\circ - \mu_\mathrm{A}{}^\circ - 2\mu_\mathrm{B}{}^\circ) = -\Delta_r G^\circ \qquad (10)$$

$$K = \left(\frac{a_\mathrm{C}{}^3 a_\mathrm{D}{}^4}{a_\mathrm{A} a_\mathrm{B}{}^2}\right)_{\text{equil}} \qquad (11)$$

$$RT \ln K = -\Delta_r G^\circ \qquad (12)$$

Since the activities can be correlated with partial pressures or concentrations through fugacity coefficients or activity coefficients, this thermodynamic approach eliminates the uncertainties otherwise associated with equilibrium calculations that are based on the law of mass action. *See* CHEMICAL EQUILIBRIUM; FUGACITY.

The prediction of equilibrium constant then requires the calculation of $\Delta_r G^\circ$ for the reaction. The so-called third-law method involves calculation for the reaction at 25°C (77°F) of the value of $\Delta_r H^\circ$, the standard enthalpy of reaction, from tabulated standard enthalpy of formation data and of $\Delta_r S^\circ$ from tabulated third-law entropies. These are combined in the sense of $\Delta_r G^\circ = \Delta_r H^\circ - T\Delta_r S^\circ$ to permit calculation of the equilibrium constant for 25°C (77°F).

This in turn is used for evaluation of the integration constant in the integration of the relation in Eq. (13).

$$\frac{d\ln K}{dT} = \frac{\Delta_r H^\circ}{RT^2} \qquad (13)$$

The integration requires expression of $\Delta_r H^\circ$ as a function of temperature, which necessitates a knowledge of the heat capacities $C^\circ_{P(i)}$ for the various reactants over the temperature range involved.

Alternatively, if values of the free-energy function $(G^\circ - H^\circ{}_{298\mathrm{K}}/T)$ are available, either from experimental measurement or from statistical thermodynamical computations, they can be combined with the standard heat of reaction at 25°C (77°F) to give the desired result, Eq. (14).

$$\Delta_r G^\circ = \Delta_r \left(\frac{G^\circ H^\circ{}_{298K}}{T}\right) + \frac{\Delta_r H^\circ{}_{298K}}{T} \qquad (14)$$

See ENTROPY; HEAT CAPACITY; THERMOCHEMISTRY.

Paul Bender

Bibliography. P. W. Atkins, *Physical Chemistry*, 1998; G. Carrington, *Basic Thermodynamics*, 4th ed., 1994; K. G. Denbigh, *Principles of Chemical Equilibrium*, 4th ed., 1981; K. S. Pitzer and L. Brewer, *Thermodynamics*, rev. ed., 1961; P. A. Rock, *Chemical Thermodynamics*, 1983; K. Wark, *Thermodynamics*, 5th ed., 1988.

Free fall

The accelerated motion toward the center of the Earth of a body acted on by the Earth's gravitational attraction and by no other force. If a body falls freely from rest near the surface of the Earth, it gains a velocity of approximately 9.8 m/s every second. Thus, the acceleration of gravity g equals 9.8 m/s^2 or 32.16 ft/s^2. This acceleration is independent of the mass or nature of the falling body. For short distances of free fall, the value of g may be considered constant. After t seconds the velocity v_t of a body falling from rest near the Earth is given by Eq. (1).

$$v_t = gt \qquad (1)$$

If a falling body has an initial constant velocity in any direction, it retains that velocity if no other forces are present. If other forces are present, they may change the observed direction and rate of fall of the body, but they do not change the Earth's gravitational pull; therefore a body may still be thought of as freely "falling" even though the resultant observed motion is upward. In the case of an initial downward velocity v_0, the downward velocity v_t after t seconds is given by Eq. (2). In the case of a body projected

$$v_t\ (\text{down}) = v_0 + gt \qquad (2)$$

upward with an initial velocity v_0, the resultant upward velocity after $t$seconds is given by Eq. (3), and

$$v_t\ (\text{up}) = v_0 - gt \qquad (3)$$

the body loses upward velocity; that is, it is accelerated downward at the rate of free fall, just as it would be if it were falling. When its resultant upward velocity has been reduced to zero, the rising body has reached its maximum height.

For a body falling a very large distance from the Earth, the acceleration of gravity can no longer be considered constant. According to Newton's law of gravitation, the force between any two bodies varies inversely with the square of the distance between them; therefore with increasing distance between any body and the Earth, the acceleration of the body toward the Earth decreases rapidly. The final velocity v_f, attained when a body falls freely from a very large distance h, involves integration of the force equation as a function of distance giving as a result Eq. (4), where R is the radius of the Earth.

$$v_f = \sqrt{\frac{2ghR^2}{R^2 + Rh}} \qquad (4)$$

In free fall from an infinite distance to the surface of the Earth, Eq. (4) reduces to $v_f = \sqrt{2gR}$, which gives a numerical value of 11.3 km/s or 7 mi/s. This is consequently the "escape velocity," the initial upward velocity for a rising body to completely overcome the Earth's attraction.

Because of the independent action of the forces involved, a ball thrown horizontally or a projectile fired horizontally with velocity v will be accelerated downward at the same rate as a body falling from rest, regardless of the horizontal motion (**Fig. 1**). If an observer could travel with the same initial horizontal velocity as the ball or the projectile, the observer would remain above it and in the observer's frame of reference the observed motion would be only that of free fall downward.

At a sufficiently large horizontal velocity, a projectile would fall from the horizontal only at the same rate that the surface of the Earth curves away beneath it (**Fig. 2**). The projectile would thus remain at the same elevation above the Earth and in effect become an earth satellite. The critical horizontal velocity v_h at which this would occur is that for which the required inward acceleration v^2/r for a body to

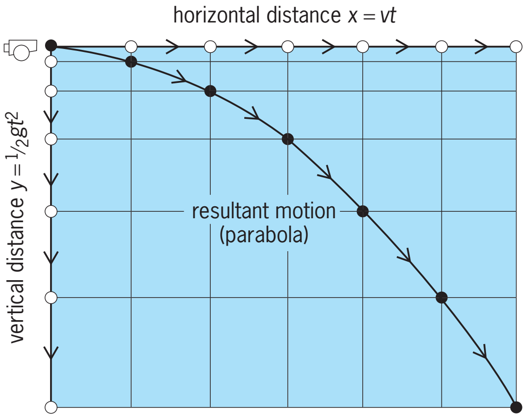


Fig. 1. Path of projectile fired horizontally, illustrating independence of horizontal and vertical components of motion.

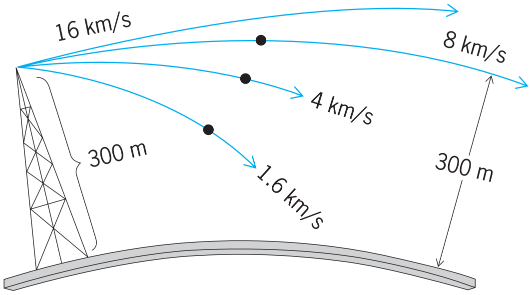


Fig. 2. Paths of projectiles fired horizontally at high velocities over the curved surface of the Earth. A projectile fired horizontally with a velocity of 8 km/s (4.8 mi/s) would circle the Earth like a satellite at constant altitude. 1 m = 3.3 ft; 1 km = 0.6 mi.

be in uniform circular motion at velocity v about a circle of radius r is exactly supplied by the free-fall acceleration of gravity g at that elevation. Thus, for the special case of horizontal motion near the Earth's surface, where $v_b = v$ and $r = R$ (the radius of the Earth), the required orbital velocity for the projectile to go into satellite motion is given by Eq. (5), and is

$$v_b = \sqrt{gR} \tag{5}$$

close to 8 km/s or 5 mi/s. At a higher velocity the satellite would rise to a higher orbit.

It was Newton's perception that for the Moon to be a satellite of the Earth the Moon must be in continual free fall toward the Earth which led him to propose his universal law of gravitation. *See* EARTH, GRAVITY FIELD OF; GRAVITATION.

Rogers D. Rusk

Free radical

Any molecule or atom that possesses one unpaired electron. This definition does not include transition- metal ions or molecules with more than one unpaired electron, such as dioxygen (O_2). Most chemists accept this definition; however, some spectroscopists employ a looser definition, in which they define any transient species (atom, molecule, or ion) in the gas phase to be a free radical.

Free radicals can be very reactive chemically (for example, the methyl radical, $CH_3\cdot$) or they can be very stable entities (for example, nitric oxide, NO). Often a free radical is designated by a dot (·). The modifier "free" probably arose because organic chemists often used to refer to a substituent (for example, hydroxyl or carboxyl) as a radical. However, this usage of the term radical is no longer prevalent, and increasingly the terms radical and free radical are used synonymously.

The term biradical is used for a molecule that contains two unpaired electrons so that the two are clearly confined to different parts of the molecule and do not interact significantly. That is, a biradical corresponds to two free radicals that are chemically bonded to each other. If the two unpaired electrons interact significantly, the entity is referred to as a triplet state. *See* TRIPLET STATE.

Free radicals were frequently, and incorrectly, postulated throughout the nineteenth century. Avogadro's hypothesis was not taken too seriously by the early organic chemists, and substances such as C_2H_6 and C_4H_{10} were frequently described as CH_3 and C_2H_5, respectively. However, by the end of that century this situation had been cleared up, and the "impossibility" of the independent existence of free radicals appeared to be well established. This situation was completely upset by Moses Gomberg's discovery of the triphenyl methyl radical during an attempt to synthesize hexaphenylethane. He found that the solution, which was expected to be colorless, was in fact yellow. Furthermore, on heating, the yellow color intensified. Also, molecular-weight determinations gave values lower than expected, and the new species was very reactive toward O_2, I_2, and NO at room temperature. This evidence provided strong support for the hypothesis that hexaphenylethane in solution is dissociated according to reaction (1).

$$(C_6H_5)_3C\text{—}C(C_6H_5)_3 \rightleftharpoons 2(C_6H_5)_3C\cdot \tag{1}$$

In spite of this evidence, Gomberg's claim met with years of doubt and denial. Since then, many similar radicals have been discovered, and such entities are now freely postulated in organic reaction mechanisms. *See* ORGANIC REACTION MECHANISM.

In 1929, the first evidence for very reactive free radicals was obtained by F. Paneth. He observed that when tetramethyllead [$Pb(CH_3)_4$] vapor, with a carrier gas, was passed very rapidly over a heated tube, a lead mirror was formed some distance downstream from the point of heating, and ethane was detected as the principal gaseous product. If the tube was then heated upstream of the mirror with tetramethyllead passing through, the old mirror was observed to disappear. This was interpreted in terms of indirect evidence that methyl radicals ($CH_3\cdot$) were present as transient free radicals in the gas phase.

Free radicals can be grouped into three major classes: atoms (for example, H·, F·, and Cl·), inorganic radicals (for example, OH·, CN·, NO·, and $ClO_3\cdot$), and organic radicals [for example, methyl ($CH_3\cdot$), allyl ($H_2C\text{—}CH\cdot\text{—}CH_2$), and phenyl ($C_6H_5\cdot$)]. Such radicals are of great importance because they have been shown to appear as intermediates in both thermal and photochemical reactions. Free radicals are also known to initiate and propagate polymerization and combustion reactions. *See* COMBUSTION; PHOTOCHEMISTRY; POLYMERIZATION; REACTIVE INTERMEDIATES.

Production. Most often, free radicals are formed by the homolytic rupture of a bond in a stable molecule with the production of two fragments, each with an unpaired electron. The resulting free radicals may participate in further reactions or may combine to reform the original compound. In the gas phase, equilibria such as reaction (2) may be established,

$$R_1\text{—}R_2 \rightleftharpoons R_1\cdot + R_2\cdot \tag{2}$$

especially at elevated temperatures. Because in most cases recombination [the leftward reaction of the equilibrium shown as reaction (2)] occurs at nearly every collision of R_1 and R_2, the composition of the reaction mixture under ordinary circumstances indicates only a minute amount of decomposition into radicals. Even though radical partial pressures may be less than 10^{-6} torr (1.33×10^{-4} pascal) and their lifetimes less than 1 millisecond, these radicals do play an important role in reaction kinetics. The transitory existence of such radicals has been confirmed by spectroscopic study. *See* CHEMICAL DYNAMICS.

Free radicals may be generated by a wide variety of methods, such as thermal decomposition, electric discharge, microwave discharge, photochemical decomposition, electrolysis at an electrode, rapid mixing of two reactants, and radiolysis with gamma-ray, x-ray, or electron-beam irradiation. In thermal methods, a stable molecule is decomposed at an elevated temperature. In exceptional circumstances, the dissociation into radicals at equilibrium may be considerable. For example, hydrogen atoms may be produced by heating hydrogen gas (H_2) to a very high temperature, as in reaction (3). At 1900 K (2960°F),

$$H_2 \rightleftharpoons 2H\cdot \qquad (3)$$

this equilibrium corresponds to 1% dissociation into atoms when the pressure is 1 atm (100 kilopascals). In a few cases, substances in solution are considerably dissociated into radicals at room temperature. It is then possible to obtain radicals of apparent long life in high concentrations. Examples are hexaphenylethane [reaction (1)], which in a 2–3% solution in benzene at 5°C (41°F) is about 3% dissociated into triphenylmethyl radicals; hexa(*p*-biphenyl) ethane [$(C_6H_5{-}C_6H_4)_3C{-}C(C_6H_5{-}C_6H_4)_3$] is virtually 100% dissociated under similar circumstances. However, this type of equilibrium is the exception rather than the rule, as most thermal decomposition reactions are irreversible under the conditions of the reaction. Most gaseous organic substances decompose wholly or in part by a mechanism involving an initial split into radicals. *See* PYROLYSIS.

Free radicals may also be produced by passing a gas through an electrical discharge at high speed. Atoms, such as hydrogen, oxygen, or nitrogen, are often produced in this manner and then subjected to further chemical reactions which often produce other radicals. A gentler method is to use a microwave discharge.

Photochemical methods of producing free radicals are very common and have been studied extensively. When subjected to ultraviolet light with wavelengths less than 200 nanometers, most gaseous organic molecules decompose photochemically into free radicals. Certain classes of molecules (for example, carbonyls) will react photochemically with ultraviolet light from 200 to 350 nm wavelength, while a few molecules, such as NO_2 and Cl_2, will decompose with visible light. When photolyzed in the range 250–310 nm, acetone splits, as in reaction (4). This

$$CH_3COCH_3 + h\nu \longrightarrow CH_3\cdot + CH_3CO\cdot \qquad (4)$$

reaction has been one of the most frequently used sources of methyl and acetyl radicals in gas-phase photochemistry.

Chlorine decomposes on irradiation with visible light and near-ultraviolet light to give chlorine atoms [reaction (5)]. This is often the starting step in the

$$Cl_2 + h\nu \longrightarrow 2Cl\cdot \qquad (5)$$

chlorination of hydrocarbons, reactions which are particularly important in the synthesis of herbicides and pesticides.

The free radical $NO_2\cdot$ decomposes via reaction (6)

$$NO_2\cdot + h\nu \longrightarrow NO\cdot + O \qquad (6)$$

with visible and near-ultraviolet light. This reaction is of considerable importance in the chemistry of air pollution.

Hydrogen peroxide dissociates photochemically with ultraviolet light of wavelengths less than 300 nm according to reaction (7). This reaction is used industrially

$$H_2O_2 \xrightarrow{h\nu} 2\cdot OH \qquad (7)$$

in advanced oxidation processes, which utilize powerful ultraviolet lamps (up to 30 kW each) to destroy organic pollutants in contaminated ground and industrial process waters. *See* AIR POLLUTION; ATMOSPHERIC CHEMISTRY.

A photochemical production of free radicals via an electron transfer reaction from a chlorophyll entity to a quinone is one of the primary photochemical reactions in photosynthesis as carried out by green plants, algae, and certain photosynthesis bacteria. *See* ELECTRON-TRANSFER REACTION; PHOTOCHEMISTRY; PHOTOSYNTHESIS.

The field of phototherapy has shown considerable promise in cancer therapy. Here the patient is injected with a dye (for example, hematoporphyrine) that has a marked affinity for tumor cells. After a couple of days, when most of the dye has concentrated in the tumor, the patient is irradiated with red light, usually with a laser. This light is absorbed by the dye, raising it to an excited state. In this excited state, the dye generates free radicals that attack the tumor cells and destroy them. *See* ONCOLOGY.

Electrolysis often produces radicals at one of the electrodes. For instance, the electrolysis of anthracene in dimethylformamide, in the absence of oxygen and water, leads to the generation of the anthracene negative radical ion at the cathode and the positive radical ion at the anode. When an anthracene anion and cation diffuse together, there is enough energy to generate the excited anthracene singlet state. This is the basis of electroluminescence. *See* ELECTROLUMINESCENCE; ELECTROLYSIS.

Rapid mixing of two suitable reactants can lead to moderate steady-state concentrations of unstable radical intermediates. For example, rapid mixing of titanium trichloride ($TiCl_3$), and of hydrogen peroxide (H_2O_2), results in the formation of the $OH\cdot$ radical, presumably by reaction (8).

$$Ti^{3+} + H_2O_2 \longrightarrow Ti^{4+} + OH\cdot + OH^- \qquad (8)$$

High-energy radiation such as gamma rays and x-rays, and high-energy particles given off by radioactive nuclei and high-energy electron beams can cause extensive disruption of molecules leading to production of radicals and ions in a process known as radiolysis. For example, the radiolysis of water produces hydrogen atoms (H·), hydroxyl radicals (OH·), and hydrated electrons ($e^{-}_{\mathrm{H_2}}$O). Each of these radical intermediates can initiate many possible secondary reactions with solutes. Radiolysis is an important technique in cancer therapy. Pulsed radiolysis, as generated by very short pulses in an electron beam, has proven to be a very important technique in the study of the mechanism of radiolytic reactions. *See* CHARGED PARTICLE BEAMS; MOLECULAR MECHANICS.

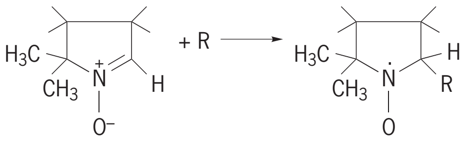
Detection and estimation. The earliest methods of detection involved the chemical properties of the radicals, such as the Paneth method for the detection of methyl radicals. Later and more reliable methods used absorption spectroscopy, mass spectrometry, and especially electron paramagnetic resonance spectroscopy.

Certain chemicals have proved effective as radical traps. For example, if iodine is added to a gas-phase photochemical system, rapid reactions (9) and (10) occur. Under appropriate conditions, all radicals

$$\mathrm{R\cdot} + \mathrm{I_2} \longrightarrow \mathrm{RI} + \mathrm{I\cdot} \qquad (9)$$

$$\mathrm{R\cdot} + \mathrm{I\cdot} \longrightarrow \mathrm{RI} \qquad (10)$$

may be removed from the system and trapped as iodides. In solution, nitrones such as 5,5′-dimethyl-1-pyrroline-1-oxide (DMPO) have proved to be effective spin traps via reaction (11). The resulting



(11)

spin adduct is usually a stable nitroxide free radical which may be detected by electron paramagnetic resonance spectroscopy. In this manner, both quantitative and qualitative information can be obtained about the nature of the reactive radical R·. For example, hydroxyl radicals (OH·) have been detected by this method in the ultraviolet photolysis of aqueous suspensions of the anatase crystal form of titanium dioxide (TiO_2). Photolysis of such suspensions has been shown to catalyze the photodegradation of chlorohydrocarbon pollutants. *See* PHOTODEGRADATION.

Absorption spectroscopy is a simple method of detecting radicals in the gas phase and has been used extensively for many years. For example, the following transient radicals (under the spectroscopists' definition) have been identified spectroscopically: OH, NH, NH_2, CN, CF_2, C_2, CH, CHO, C_3, CH_3O, C_2H_5O, NCO, NCS, HNO, PH_2, and CH_3; many others have been identified with varying degrees of certainty. The knowledge of the absorption wavelengths of these radicals has permitted astronomers to detect many of these entities in interstellar space. The flash-photolysis technique, by which large amounts of light energy may be absorbed in a short period of time, has greatly advanced the detection of radicals by absorption spectroscopy and their involvement in radical mechanisms.

Another popular technique for radical detection is matrix isolation. Here the radicals are generated (for example, by photolysis) in the gas phase in an inert gas such as argon. Immediately after generation, the gas mixture is frozen onto a finger, a protusion into the gas stream that is maintained at a very low temperature (less than 20 K or -423°F). Alternatively, molecules may be isolated in a frozen argon matrix and then photolyzed to generate the radicals. Once isolated in the argon matrix by this method, the radical intermediate may be studied at will by using ultraviolet, visible, or infrared spectroscopy. *See* LUMINESCENCE ANALYSIS; MATRIX ISOLATION; SPECTROSCOPY.

An ingenious method for the detection of gas-phase radicals has been developed using mass spectrometry. A rapid-flow system is used in which radicals are produced thermally or protochemically. The gas stream is sampled through a leak into the mass spectrometer. An electron energy is selected so that radicals are ionized, but stable molecules are not dissociated. Any radical ions detected in the mass spectrometer must come from radicals already present in the gas mixture, and a direct and unequivocal detection of radical intermediates is possible. Under favorable circumstances, the method can be used for the quantitative estimation of radical concentrations. This technique has proved to be particularly valuable in the study of combustion reactions. *See* MASS SPECTROMETRY.

Electron paramagnetic resonance spectroscopy is by far the best technique available for the detection and characterization of free radicals in solids, liquids, or gases. The technique is very sensitive (10^{-9} mole per liter free radicals can be detected under favorable conditions) and, furthermore, the technique is specific to molecules that contain one or more unpaired electrons. Generally, each radical has a characteristic electron paramagnetic resonance spectrum which can be used both for identification and for deducing structural information about the radical. For example, the methyl radical $CH_3\cdot$ was shown to have a planar structure by analyzing its electron paramagnetic resonance spectrum, whereas the corresponding $CF_3\cdot$ radical was shown to have a pyramidal structure. Analysis of the electron paramagnetic resonance spectra of radicals such as the anthracene negative ion yields information concerning the distribution of the unpaired electron over the molecule. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY.

Free-radical mechanisms. In well-investigated reactions, it is frequently possible to establish that the reaction proceeds via a multistep mechanism involving free-radical intermediates. For example, a 50:50 mixture of chlorine and hydrogen will explode on being subjected to a flash from a photoflash unit. This reaction is thought to proceed via the mechanism outlined in reactions (12)–(16). This is an example of

$$\mathrm{Cl_2} + h\nu \longrightarrow \mathrm{2Cl\cdot} \qquad (12)$$

$$\mathrm{Cl\cdot} + \mathrm{H_2} \longrightarrow \mathrm{HCl} + \mathrm{H\cdot} \qquad (13)$$

$$\mathrm{H\cdot} + \mathrm{Cl_2} \longrightarrow \mathrm{HCl} + \mathrm{Cl\cdot} \qquad (14)$$

$$\mathrm{H\cdot} + \mathrm{HCl} \longrightarrow \mathrm{H_2} + \mathrm{Cl\cdot} \qquad (15)$$

$$\mathrm{Cl\cdot} + \mathrm{X} \longrightarrow \text{nonradical products} \qquad (16)$$

a radical chain reaction since reactions (13) and (14) can cycle many times generating the product HCl before the chain is terminated by reaction (16) [X in reaction (16) is any substance which can remove Cl· radicals from the system]. The system explodes because reactions (13) and (14) are extremely fast and the reaction generates a considerable amount of heat, which further accelerates the reaction. Many thermal decomposition reactions are also thought to proceed via free-radical chain reactions. A typical example is the thermal decomposition of ethane, reactions (17)–(23). General mechanisms of this type,

$$\mathrm{C_2H_6} \longrightarrow \mathrm{2CH_3\cdot} \qquad (17)$$

$$\mathrm{CH_3\cdot} + \mathrm{C_2H_6} \longrightarrow \mathrm{CH_4} + \mathrm{C_2H_5\cdot} \qquad (18)$$

$$\mathrm{C_2H_5\cdot} \longrightarrow \mathrm{C_2H_4} + \mathrm{H\cdot} \qquad (19)$$

$$\mathrm{H\cdot} + \mathrm{C_2H_6} \longrightarrow \mathrm{C_2H_5\cdot} + \mathrm{H_2} \qquad (20)$$

$$\mathrm{H\cdot} + \mathrm{C_2H_5\cdot} \longrightarrow \mathrm{C_2H_6} \qquad (21)$$

$$\mathrm{2C_2H_5\cdot} \longrightarrow \mathrm{C_4H_{10}} \qquad (22)$$

$$\mathrm{2C_2H_5\cdot} \longrightarrow \mathrm{C_2H_4} + \mathrm{C_2H_6} \qquad (23)$$

the Rice mechanisms, have been well established in many cases by detailed investigation of the kinetic behavior. With the aid of computer simulation, it is possible to predict the behavior of very complex reactions over a wide range of conditions. This approach has been particularly valuable in the study of gas-phase chemistry in the upper atmosphere and in polluted air. *See* CHAIN REACTION (CHEMISTRY); COMPUTATIONAL CHEMISTRY. James R. Bolton

Bibliography. A. Carrington and B. Thrush (eds.), *The Spectroscopy of Molecular Ions*, 1988; B. Halliwell, *Free Radicals in Biology and Medicine*, 3d ed., 1999; H. Kurreck, *Electron Nuclear Double Resonance of Radicals in Solution*, 1988; W. A. Pryor, *Frontiers of Free Radical Chemistry*, 1980; J. A. Weil, J. R. Bolton, and J. E. Wertz *Electron Paramagnetic Resonance*, 1994.

Frequency (wave motion)

The number of times which sound pressure, electric field, or other quantities specifying a wave vary (or oscillate) from their equilibrium value through a complete cycle in a specified time. The most common unit of frequency is the hertz; 1 Hz is equal to 1 cycle per second. In one cycle there is a positive variation from equilibrium, a return to equilibrium, then a negative variation, and return to equilibrium. This relationship is often described in terms of the sine wave, and the frequency referred to is that of an equivalent sine-wave variation in the parameter under discussion. *See* SINE WAVE.

The time between consecutive variations or oscillations of the wave is the period T of the wave. The frequency f is the reciprocal of the period:

$$f = 1/T$$

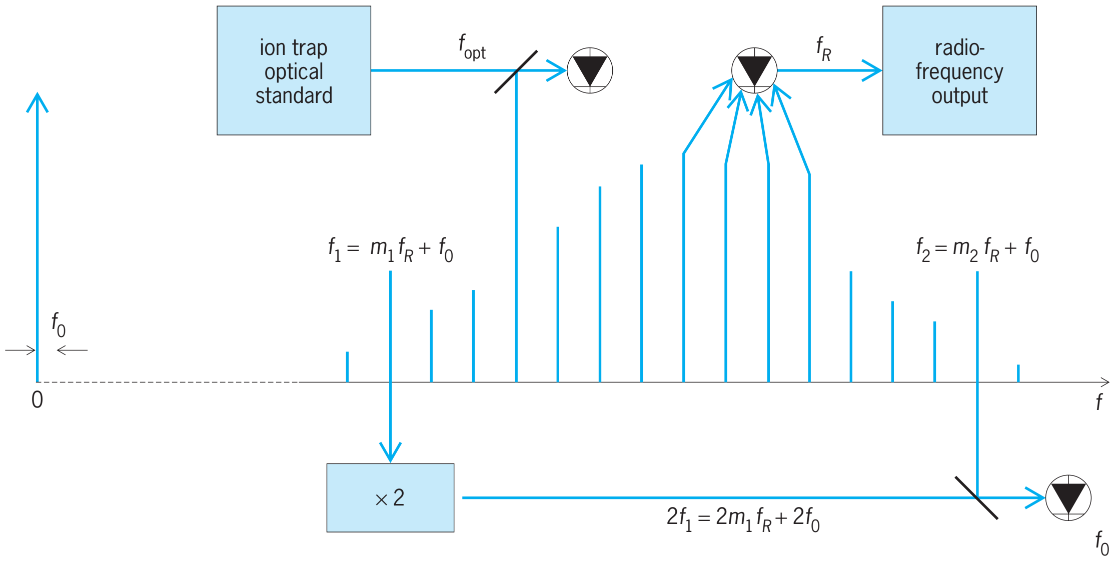
Frequency is a convenient means for describing the various ranges of interest in wave motion. For example, audible sound is between approximately 20 and 20,000 Hz. Infrasonic frequencies are below approximately 20 Hz; sound having frequencies above the audible is termed ultrasonic. *See* INFRASOUND; SOUND; ULTRASONICS.

Ultrasonic waves having frequencies as high as 200 gigahertz have been generated and detected in solids and have been used to generate images ($1\ \mathrm{GHz} = 10^9\ \mathrm{Hz}$). Typical approximate frequency ranges for electromagnetic waves are: AM radio in the United States, 550 to 1700 kHz; FM radio, 88 to 108 MHz; Global Positioning System (GPS) microwaves, 1.2276 GHz and 1.57542 GHz; visible light, 4×10^{14} to 7.5×10^{14} Hz. *See* ELECTROMAGNETIC RADIATION; FREQUENCY MEASUREMENT; WAVE MOTION. William J. Galloway; Philip L. Marston

Frequency comb

A device that facilitates and greatly simplifies the measurement of very high frequencies, particularly in the optical range from 10^{14} to 10^{15} Hz. Extremely stable and reproducible optical frequencies f_{opt} in this range are produced by trapped-atom and trapped-ion atomic clocks. Frequency-comb technology is based on the train of short-duration pulses of optical radiation emitted at a stable and measurable repetition rate f_R of 100 MHz to 1 GHz by a femtosecond laser. Such a pulse train can be equivalently represented as a continuously produced "comb" of frequencies. The "teeth" of the comb are discrete phase-related frequencies spaced by f_R from one another, and the frequency of the mth tooth is mf_R, where m is an integer (see **illustration**). The frequencies produced typically span a rather narrow wavelength range of 30 nm to around 800 nm, but this span can be increased to a factor of 2 by passing the pulsed light through an optical fiber made from a highly nonlinear material, enabling a comb frequency to be matched to many different optical frequency standards. *See* ATOMIC CLOCK.

The frequencies of the comb as a whole are offset by f_0 because the optical frequency of the femtosecond laser in general differs by a frequency f_0 from an exact multiple of f_R. One of the comb frequencies, $mf_R + f_0$, can be mixed with f_{opt}, and the output from the mixing diode used as feedback to stabilize this comb frequency and confer on it the full frequency stability of f_{opt}. The frequency f_0 can be measured by frequency doubling a frequency in the left-hand



Operation of the frequency comb, illustrating the self-referencing technique used to measure the offset frequency f_0. (*Hugh Klein*)

(infrared) end of the comb (by passing the light through a nonlinear optical medium) to generate a frequency $2m_1f_R + 2f_0$, and mixing this frequency with that of a tooth of frequency $m_2f_R + f_0$ from the right-hand end of the comb. If the chosen tooth has $m_2 = 2m_1$, the beat frequency is f_0 and can be measured. *See* BEAT.

Provided the integer m is known for the tooth frequency nearest to f_{opt} (from an approximate knowledge of f_{opt} and f_R), f_R can be calculated as $f_{\mathrm{opt}}/m - f_0/m$. In this way the long-term stability of optical frequency standards can be subdivided to a radio frequency f_R for extremely accurate frequency and time measurement. The fractional stability of f_R is the same as that of the optical frequency standard to which it is locked and this is now of the order of 1 in 10^{15}, equaling that of the very best cesium clocks. This stability will undoubtedly improve even further. Then it is anticipated that the SI second will be redefined in terms of an optical frequency standard. *See* FREQUENCY MEASUREMENT; PHYSICAL MEASUREMENT; TIME.

Bryan P. Kibble

Bibliography. *Proceedings of the First ESA Optical Clock Workshop*, 2005; F. Riehle, *Frequency Standards*, Wiley-VCH, 2004; Special issue: 50 years of atomic time-keeping: 1955 to 2005, *Metrologia*, 42(3):S1–S153, June 2005.

Frequency counter

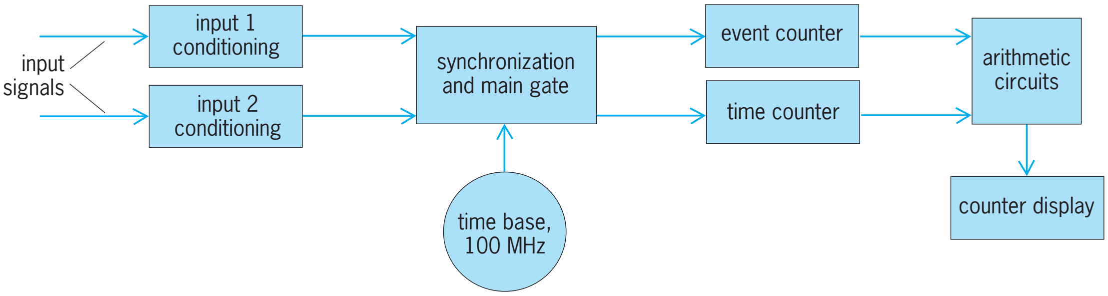
An electronic instrument used to precisely measure the frequency of an input signal. Frequency counters are commonly used in laboratories, factories, and field environments to provide direct frequency measurements of various devices. The most common applications for frequency counters are measurement and characterization of oscillator and transmitter frequencies. *See* OSCILLATOR; RADIO TRANSMITTER.

There are several classes of frequency counters. Basic frequency counters provide measurement of frequency only. Universal counter-timers are two-channel instruments that provide measurement of frequency, period, phase, totalize (the total number of pulses generated by some type of event over the duration of an experiment), ratio (of frequencies on two channels), and time intervals such as pulse width or rise time. Microwave counters are an extension of basic frequency counters offering coverage of microwave frequency ranges to 40 GHz and beyond. *See* MICROWAVE.

Architectures. The three main architectures are conventional counting, reciprocal counting, and continuous counting. Conventional counting is the oldest and simplest but has the lowest performance and the least measurement flexibility. A conventional counter uses a simple register to count each cycle of the input signal during a 1-s measurement gate time. Since the count register contains a direct count of the number of cycles in one second (hertz), its contents can be sent to a digital display as the measurement result with no further processing.

Reciprocal counting is the most common architecture (see **illus.**). It provides improved performance and flexibility over conventional counting. The main gate is set by the user and determines the nominal time over which the measurement is to be made (measurement gate time). The input signal is applied to the input conditioning circuits and routed to the event counter. The event counter simply counts the exact number of cycles of the input signal during the measurement gate time. The reciprocal technique synchronizes the opening and closing of the main gate to the input signal. This provides increased precision but requires that the actual measurement gate time also be measured precisely. This is the purpose of the time counter. The time counter counts the number of cycles of the internal time base during the measurement gate time.

At the end of the measurement gate time, the values in the time and event counters are read by the arithmetic circuits. The arithmetic circuits then



Block diagram of a reciprocal frequency counter.

compute frequency as the number of cycles per unit time. The event counter represents the number of cycles of the input signal counted, and the exact measurement time is given by the number of counts of the precision time base multiplied by the time-base period. The result is then converted to hertz and sent to the counter display.

Continuous counting is based on the reciprocal technique, but employs high-speed digital circuits to continuously sample the contents of the count registers. These continuous samples can be digitally processed to provide improved resolution to as many as 12 digits in 1 s of measurement time. Continuous counting is also employed in frequency and time-interval analyzers, which provide graphical displays of continuous frequency versus time.

Range extension. The typical upper range for a direct frequency counter is 200 MHz. Prescalers are often used to extend the range of frequency counters to 3 or 5 GHz. A prescaler effectively divides the input signal frequency down into the range of the digital counter. The prescaler is placed after the input conditioning section (see illus.). Counting is then performed as described above with the prescale factor multiplied back in by the arithmetic circuits to display the correct answer.

Harmonic heterodyne converters are typically used to extend the measurement range to 40 GHz and beyond. In its simplest form, the microwave input signal is mixed in a sampler with a frequency generated by a synthesizer to provide an intermediate frequency that is in the direct measurement range of the digital counter. Modern microwave frequency counters integrate the harmonic heterodyne converter into their design to automatically downconvert and measure microwave frequencies. *See* FREQUENCY MEASUREMENT; HETERODYNE PRINCIPLE.

Barry Dropping

Bibliography. C. F. Coombs, Jr., and G. D. Sasaki, *Electronic Instrument Handbook*, 3d ed., 1999.

Frequency divider

An electronic circuit that produces an output signal at a frequency that is an integral submultiple of the frequency of the input signal.

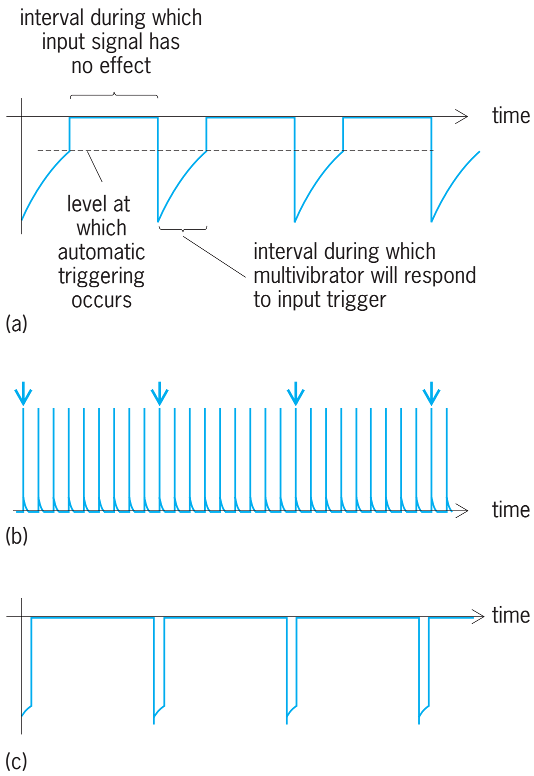
Several information-processing and transmitting techniques require frequency division. In television, for example, it is essential to maintain a precise relationship between the horizontal-scanning frequency and the vertical-scanning frequency. Frequency division can be easily accomplished in two ways, digital division and division by triggering a subharmonic.

Digital division. Many circuits are available to count pulses. A bistable or flip-flop circuit produces one output pulse for every two input pulses. By cascading successive flip-flops, any desired degree of division can be obtained. Division by powers of 2 can be achieved simply by monitoring the output of the proper stage in the cascade. Division by other numbers can be achieved by gating to obtain the proper set of flip-flop conditions. Division by 10, or powers of 10, is readily obtained by using decade rings. *See* DIGITAL COUNTER; GATE CIRCUIT.

Reduction of the input to pulses can be accomplished when needed by amplifying, clipping, and differentiating the original signal. To provide an output that is a desired shape, the signal from the frequency divider can be used to synchronize a multivibrator to obtain a square wave and additional power. The square wave can be fed into a tuned circuit or filtered to obtain a sine wave, if that is needed. In many applications the input and output signals need no alteration. *See* WAVE-SHAPING CIRCUITS.

Subharmonic triggering. Any circuit which has a characteristic resonance responds to certain types of input energy by ringing, that is, by going through one or more cycles of electrical activity caused by the nature of the circuit rather than by the nature of the input. This characteristic can be used to accomplish frequency division, provided the input frequency does not vary over any extensive frequency range. A triggered multivibrator can be used in this manner. Either a monostable or a synchronized astable type will serve. (An astable type can be used because, while it is normally free-running with no stable state, it can be triggered to provide synchronized pulse rates slightly faster than it would otherwise generate.)

The component values of the frequency-dividing multivibrator are selected so that triggering cannot occur during a fixed number of cycles of the input signal. In operation, each time the multivibrator returns to the state where it is sensitive to an input trigger, the next cycle of the input signal initiates a new output pulse. Hence a new cycle of the multivibrator is initiated every *n* input cycles. The frequency of the resulting output signal is therefore an exact submultiple of the frequency of the input signal. Because the input may go through a large number of cycles for each output cycle, greater division per stage can be realized by this method than by digital division.



Waveform of frequency-dividing astable multivibrator. (*a*) Waveform at base terminal of free-running astable multivibrator, showing how multivibrator would perform if there were no input. (*b*) Input signals. Arrows indicate triggering signals. (*c*) Output signal, synchronized to every *n*th input cycle.

The input–output relationship can be most readily perceived when the input consists of short pulses (see **illus.**), but the effect is not dependent on the shapes of the input signal.

Other resonant or tuned circuits can be induced to supply submultiple outputs in a similar manner. *See* MULTIVIBRATOR; RESONANCE (ALTERNATING-CURRENT CIRCUITS).

William W. Snow

Frequency measurement

The determination of the number of cycles of a periodically varying quantity occurring in unit time. Many physical systems demonstrate cyclic behavior; that is, one or more of their properties vary in a characteristic fashion before returning to the initial value and then repeating the cycle. Examples are the angular positions of the planets and satellites in the solar system, the pressure in a cylinder in a reciprocating engine, and the heights and fields associated with surface, acoustic, and electromagnetic waves. The duration of a single cycle, the period, may vary widely, from 10^{-27} s for the electromagnetic field associated with a cosmic gamma ray to 10^8 years for the rotation of a galaxy in space. The frequency, which is the inverse of the period, is the number of cycles, including fractions, occurring in unit time. The unit of frequency is the hertz (Hz), named after Heinrich Hertz, who investigated the nature of electromagnetic radiation. Measurement of the characteristic frequencies of a system, and their variation with time or under changing conditions, yields valuable information on its properties and behavior. Together with temperature and voltage, frequency ranks as one of the quantities most often measured in modern science and technology. *See* CELESTIAL MECHANICS; CERENKOV RADIATION; ELECTROMAGNETIC RADIATION; GAMMA-RAY ASTRONOMY; MILKY WAY GALAXY; WAVE MOTION.

Analog techniques. The measurement of an unknown frequency requires a standard producing a fixed, stable, and known frequency, and a system or technique for the comparison of the unknown frequency with the standard. In the past, a wide variety of analog techniques and material standards have been employed. An example is the use of a tuning fork to adjust a musical instrument, usually a piano. The combination of notes from the instrument and the resonating fork has an amplitude modulation at the difference between their frequencies, which is detected as a slow audible beat. Similarly, two sinusoidal electrical signals close in frequency applied to the X and Y plates of an oscilloscope produce an ellipse on the screen which rotates at the beat frequency. If this frequency is small, less than 1 Hz, it may be measured by timing the period with a stopwatch. The more complicated Lissajous figures are produced if the two frequencies are (nearly) harmonically related. *See* BEAT; LISSAJOUS FIGURES; TUNING FORK.

Analog frequency measurement techniques possessed two major disadvantages: The frequency of the standards depended upon the material properties and dimensions of critical components, which meant that they were prone to drift and affected by variations in the ambient temperature. In addition, optimum accuracy was achieved only when the unknown and standard frequencies were close or harmonically related.

Developments in electron-tube and, later, solid-state electronics improved matters. These included the quartz crystal oscillator, in which a thin slice of crystalline quartz acts as the resonant element in an electronic feedback circuit. As a result of the sharpness of the resonance and the stability of the properties of the quartz, this device provides a stable frequency in the range from 10 kHz to 100 MHz and remains the most common secondary frequency standard in use. In addition, a range of circuits were developed to generate more complex harmonic and subharmonic frequencies from a standard source. This led ultimately to the frequency synthesizer which, with an array of phase-locked loops, could be set to produce one of a very wide range of output frequencies. In use, however, it was still necessary to measure the beat or heterodyne frequency from the unknown frequency. *See* FEEDBACK CIRCUIT; FREQUENCY DIVIDER; FREQUENCY MULTIPLIER; OSCILLATOR; PHASE-LOCKED LOOPS; PIEZOELECTRICITY; QUARTZ CLOCK.

Digital techniques. Fast, inexpensive solid-state digital circuits have replaced analog frequency measurement techniques and many of their associated standards. The underlying principle of the digital

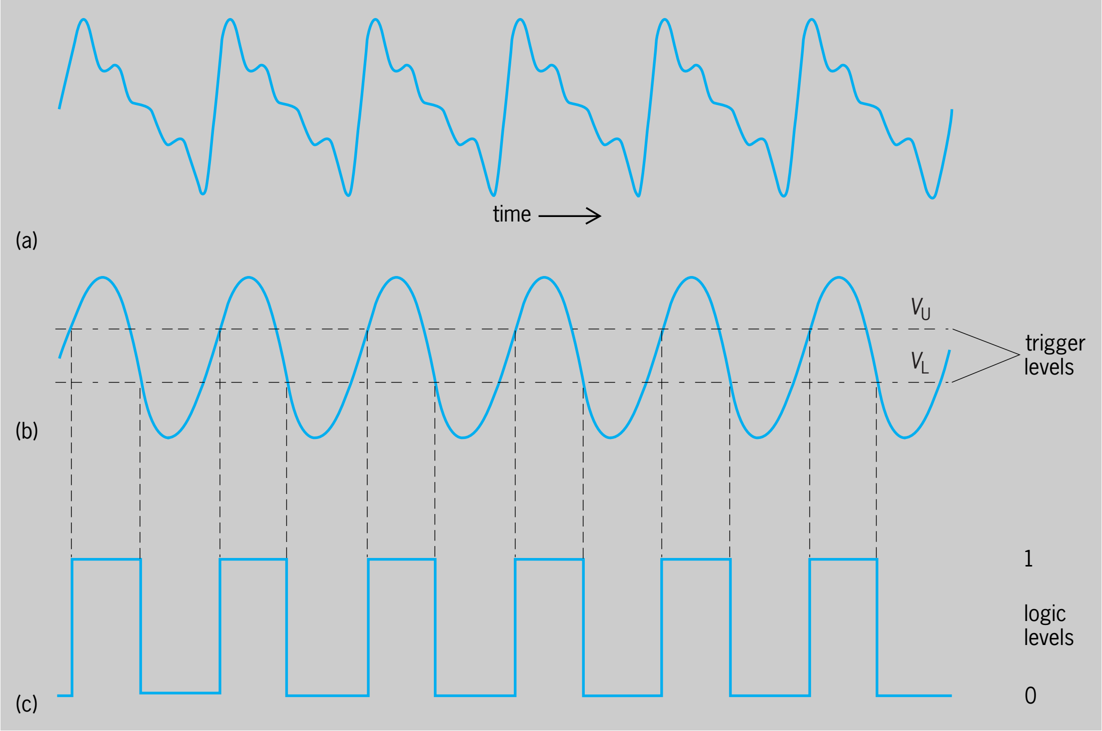


Fig. 1. Waveforms in frequency measurement. (*a*) Arbitrary cyclic variation. (*b*) Sine wave at the fundamental frequency. (*c*) Digital signal for counting.

technique is simple: the electrical signal (**Fig. 1***a*) from the sensor or transducer observing the physical system under test generally contains, from Fourier analysis, the fundamental frequency (Fig. 1*b*) and components at integral harmonics of this frequency. It is filtered to select the fundamental and converted into a rectangular waveform (Fig. 1*c*), representing transitions between the binary logic levels 0 and 1. A frequency measurement then consists of counting the number of positive- or negative-going transitions between the two levels in a known time. *See* FOURIER SERIES AND TRANSFORMS.

In parallel with the production of counters capable of operating at frequencies up to around 1 GHz, frequency standards based upon selected atomic transitions rather than the properties of bulk materials have been developed. These have the advantage that the frequency produced from a particular transition is in principle universal; that is, it is largely independent of the design of the standard and the materials used in its construction, and of changes in the ambient conditions. The combination of high-speed digital counters and of very stable atomic reference sources allows a wide range of frequencies to be determined simply, inexpensively, and very accurately.

As a result, much work has been carried out on the definition and measurement of other physical quantities in terms of frequency. Clearly, time and frequency are closely related; not only are the measurement, calibration, and dissemination techniques largely interchangeable, but any frequency standard may be converted into a standard of time, that is, a clock, by adding an appropriately designed counter. The unit of time, the second, is itself defined as the duration of 9 192 631 770 cycles of the electromagnetic radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom. The primary standards of voltage and current are now also realized in terms of frequency using the superconducting Josephson effect in conjunction with the von Klitzing (quantum Hall) effect. *See* ATOMIC TIME; ELECTRICAL UNITS AND STANDARDS; RESISTANCE MEASUREMENT; TIME; VOLTAGE MEASUREMENT.

General-purpose frequency counters. In a frequency counter, sometimes referred to as a frequency meter or timer-counter (**Fig. 2**), the input signal enters a direct-coupled broadband amplifier covering the frequency range of the instrument, whose gain can be switched to accommodate different signal levels. Protective elements may be placed at the input to prevent damage from excessive voltage levels. The amplified signal is then converted into a rectangular waveform with a bistable logic element, for example a Schmitt trigger. When the signal rises through a preset dc voltage level (trigger level V_u, Fig. 1*b*), the circuit switches very rapidly, in a few nanoseconds or less, from the 0 to the 1 state. The trigger level can be varied to select an appropriate point on the input waveform, preferably where it is varying rapidly. The trigger circuit is designed to show some hysteresis; that is, the reverse transition from 1 to 0 takes place at a lower voltage level (trigger level V_L) to reduce the possibility that noise might produce additional and unwanted transitions. *See* DIRECT-COUPLED AMPLIFIER; HYSTERESIS; MULTIVIBRATOR.

The output from the trigger circuit is passed to a counter through an electronic gate, which is opened for the required measurement period. The gating

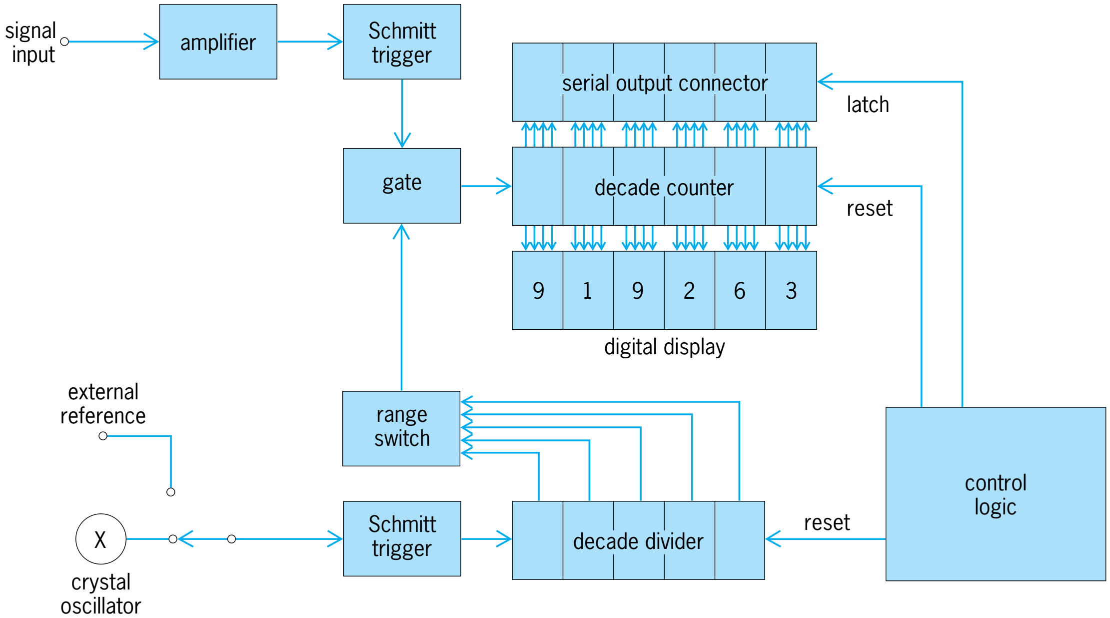


Fig. 2. Frequency-counter block diagram.

waveform is generated by dividing down a reference frequency, usually from a quartz crystal oscillator at 1, 5, or 10 MHz, with a decade divider. The waveforms from different stages of the divider may be used to provide various measurement times, for example from 1 ms to 10 s. For critical applications, an external frequency reference with high stability and accuracy may replace the internal crystal oscillator. *See* GATE CIRCUIT.

The selected transitions are counted by a decade counter, which may contain up to 10 stages, each of which consists of four solid-state bistable elements arranged to count in binary-coded decimal (BCD) code. The count is displayed on the front panel of the instrument using a light-emitting diode or fluorescent digital display. Liquid-crystal displays are sometimes found on small portable counters where battery life is important. The final count at the end of each measurement cycle is transferred (latched) into solid-state memory so that it is available for transmission to other instruments, either in parallel BCD code or after conversion to serial form. *See* DIGITAL COUNTER; ELECTRONIC DISPLAY.

The basic operations required of the control logic are to reset the counters at the start of each measurement cycle, to transfer the result to the output memory register at its conclusion, and to arrange an adjustable delay before the start of the next cycle if repeated measurements are required. Many modern frequency counters contain a microprocessor which supervises these functions, and adds others to increase the versatility of the instrument (**Fig. 3**). *See* FREQUENCY COUNTER; MICROPROCESSOR.

Measurement uncertainties. The errors in the readings displayed on the counter consist of avoidable errors arising from unwanted components in the input signal, and those inherent in the operation of the counter. Excessive noise in the signal, at a level greater than the hysteresis in the trigger circuit, gives rise to additional transitions and hence a high reading. This occurrence is usually obvious, as the number of additional counts is random and the reading varies with time. A stable reading is not, however, a guarantee that correct operation is taking place. The presence of harmonics in the signal (Fig. 1*a*) may lead to additional transitions within each cycle which add a constant proportion to the reading. On the other hand, the presence in the input of a large component at a low frequency, usually from power-line pick-up, may reduce the reading by shifting some cycles outside the trigger levels. Fortunately, each of these possibilities is easily detected by observing the input signal on an oscilloscope; it should consist mostly of the fundamental harmonic of the frequency to be measured, with the random noise and hum levels well below the hysteresis in the trigger circuit. *See* OSCILLOSCOPE.

In the absence of these gross errors, the uncertainty in the reading is the sum of two contributions. Because the counter effectively measures the ratio of

Fig. 3. General-purpose frequency counter. (*Hewlett-Packard Co.*)

the unknown frequency f to that of the reference f_0, any error Δf_0 in the latter appears proportionally in the result, as in Eq. (1), where Δf is the measurement.

$$\frac{\Delta f}{f} = -\frac{\Delta f_0}{f_0} \qquad (1)$$

The frequencies generated by quartz crystal oscillators depend upon the ambient temperature and drift steadily with time. The temperature-compensated crystal oscillator (TCXO) found in less expensive counters contains a temperature-sensitive network in the oscillator circuit to minimize thermal effects. The associated uncertainty would typically be around 10^{-6} for a year after calibration. For better performance the crystals are aged, selected for good stability, and operated in a temperature-controlled oven. A stability of 10^{-7} or better over a year may then be achieved, limited by the aging rate.

The second uncertainty arises from the resolution of the counter itself. Because the phases of the unknown and reference frequencies are in general independent, the times of the signal transitions are unrelated to the start and end of the gate waveform. As a result, the count may vary by one in the last digit. The resultant uncertainty is given by Eq. (2), where

$$\frac{\Delta f}{f} = \frac{1}{\tau f} \qquad (2)$$

τ is the gate or measurement time in seconds. For a given frequency, this uncertainty may be reduced only by increasing τ.

Measurements at low frequencies. Frequency counters often not only count the number of complete cycles within the gate period but also estimate the contribution from the partial cycles at each end. A simpler alternative uses a counter in the time interval mode, measuring the duration of a known number of cycles of the unknown frequency. Essentially, the input and reference signals (Fig. 2) are interchanged, so that the gate waveform is obtained by dividing down the unknown frequency while the reference frequency f_0 is counted for this period. As f_0 is typically between 1 and 10 MHz, the resolution uncertainty is then comparable with that of the reference frequency.

The unknown frequency is calculated from the time interval. The gate period is affected by the presence of noise in the input signal. An approximate calculation of the additional uncertainty arising from this source gives Eq. (3), where S_N is the input signal-to-noise ratio.

$$\frac{\Delta f}{f} = \frac{1}{\tau f S_N} \qquad (3)$$

This is of the same form as the resolution uncertainty, but reduced by S_N.

Microwave frequency counters. The finite transition time of counting circuits limits their operation to frequencies less than about 1 GHz. Prescalars, very fast bistable elements arranged to divide the input frequency by a factor of 2 or 4, can extend the counter operating range up to about 3 GHz. To retain the advantages of the direct counting method, a method of transforming even higher frequencies into the working range in an accurate and predictable manner is required. Two mechanisms are available for this purpose, beat-frequency and harmonic-frequency generation. If two signals at frequencies f and $f + \Delta f$ are combined, the difference or beat frequency Δf appears as an amplitude modulation of the result. For electrical signals, the beat frequency may be isolated by mixing the waveforms in a nonlinear device such as a rectifying diode and removing the higher-frequency components with a filter circuit. *See* HETERODYNE PRINCIPLE; MIXER.

Because nonlinear devices distort the signal waveform, they also produce harmonics from a clean sinusoidal input. Step recovery diodes, for example, which have a sharp nonlinearity with a transition time of a few picoseconds, can produce a comb (so called from the appearance on a spectrum analyzer) of harmonics extending across the microwave region. It is possible to combine harmonic generation and beating in a single device to produce a measurable difference frequency Δf from the standard f_0 and a much higher unknown frequency f, given by Eq. (4).

$$\Delta f = \pm[f - nf_0] \qquad (4)$$

To calculate the unknown frequency f, the appropriate sign in Eq. (4) and the harmonic number n must he determined. A microwave frequency counter based upon this principle of harmonic heterodyne down-conversion (**Fig. 4**) employs a synthesizer producing a high frequency f_S phase-locked to the internal crystal reference. The synthesizer can be programmed by a microprocessor controller to vary f_S in steps over a range, typically from 200 to 300 MHz. The frequency f_S is multiplied and mixed with the microwave frequency in the harmonic mixer forming the first stage of the counter. The measurement cycle involves reducing f_S from its maximum value until the filtered signal from the broadband intermediate-frequency amplifier contains a beat frequency Δf_1 within the range of the decade counter. This is measured, f_S is reduced by a small amount Δf_S, and a second reading Δf_2 of the beat frequency is taken. Bearing in mind that the harmonic number n must be a positive integer, Eq. (5) enables the sign and n to he determined exactly, so that the microwave frequency can be calculated and displayed.

$$n \approx \pm\left[\frac{\Delta f_2 - \Delta f_1}{\Delta f_S}\right] \qquad (5)$$

Optical frequency measurements. Above the microwave frequency range, for electromagnetic radiation in the submillimeter, infrared, and optical regions of the spectrum, frequency measurements became feasible only with the development of lasers producing narrow-bandwidth coherent radiation. Early work demonstrated that the stability of the laser frequency could be greatly improved by locking it to the centers of molecular absorption lines. By reflecting the

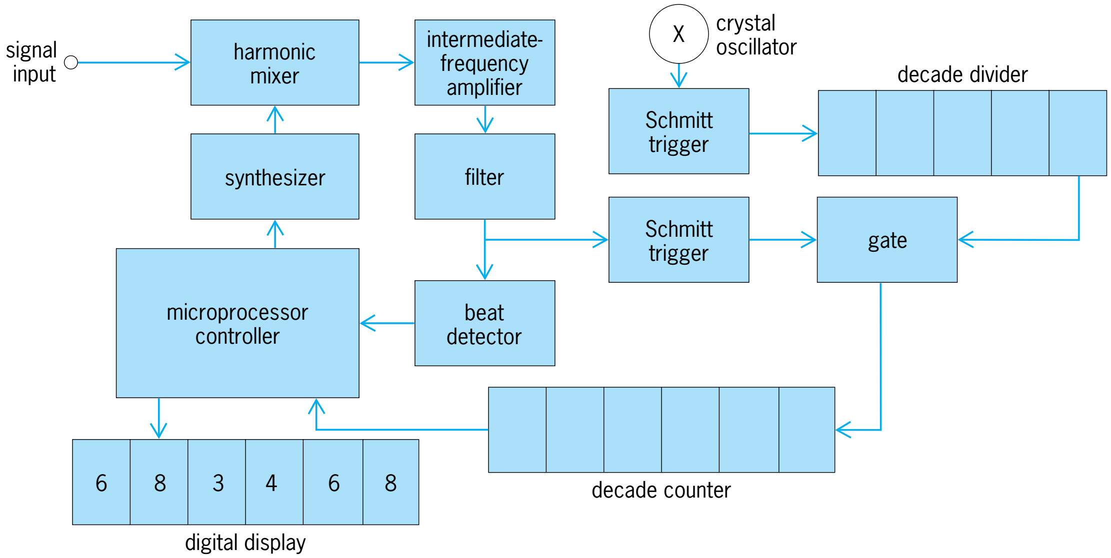


Fig. 4. Microwave frequency counter.

laser beam back through a cell containing a suitable gas at low pressure, it is possible to eliminate the line broadening from the random motion of the molecules, that is, the first-order Doppler effect. This technique of saturated absorption gives sub-Doppler linewidths of 1 MHz or less, and enables stabilities of around 1 part in 10^{12} to be achieved. *See* LASER; LASER SPECTROSCOPY.

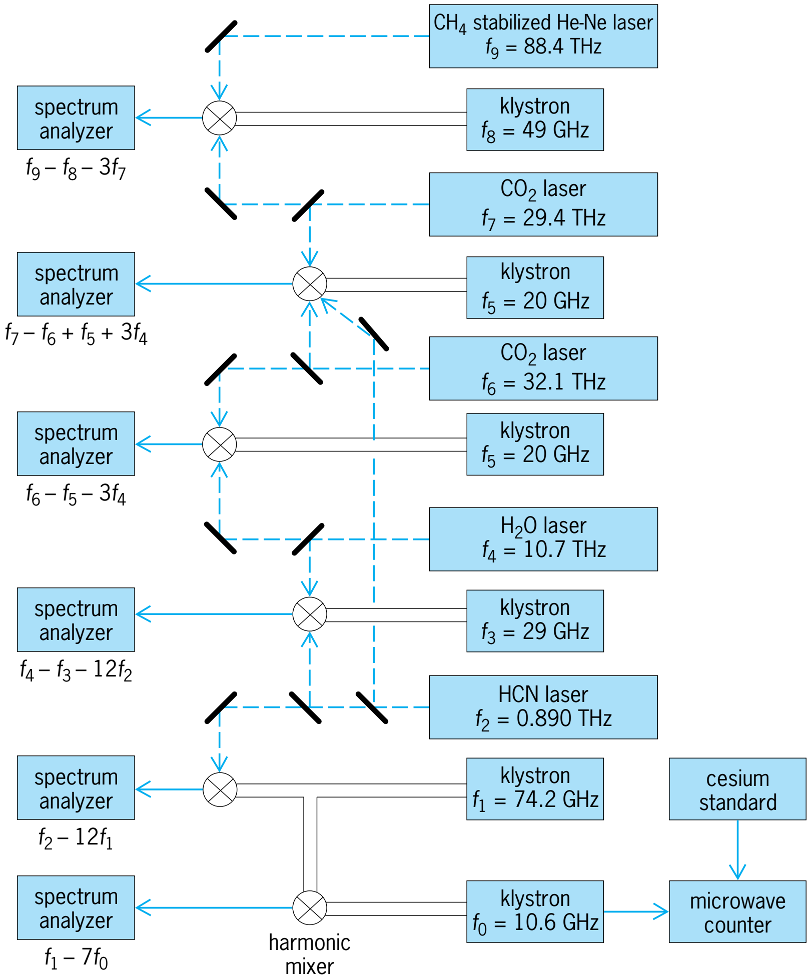


Fig. 5. Laser frequency measurement chain.

The equipment and techniques required for optical frequency measurement are complex and of value only in scientific and metrological research. A major reason for the early work was the possibility of obtaining a more accurate value for the speed of light c_0 from the frequency v and vacuum wavelength λ_0 of a stabilized laser, using the relationship $v\lambda_0 = c_0$. The first precise determination of c_0 by this method was made in 1972 through the measurement of the frequency of a methane-stabilized helium-neon laser at 88.4 THz, corresponding to a wavelength of 3.39 micrometers (**Fig. 5**). The harmonic mixers in this and later work were point-contact devices bearing a close resemblance to the "cat's whisker" detectors from the early days of radio receivers. Here the whisker also serves as an antenna to couple the laser radiation into the mixer. Although the stabilized lasers were selected so that the best frequencies between their harmonics were small, microwave radiation from a stabilized klystron was also necessary to down-convert to a radio frequency within the range (30–1500 MHz) of the spectrum analyzers which monitored and measured the beats. All frequencies were referred to that (f_0) from an X-band klystron which was calibrated against a cesium atomic beam standard. *See* KLYSTRON; SPECTRUM ANALYZER.

The value obtained for c_0, 299 792 458 m · s^{-1}, was assigned an uncertainty of 4 × 10^{-9}, which arose mainly from the difficulty of comparing the laser wavelength to that of the krypton-86 emission line then forming the standard of length. In 1983 the General Conference of Weights and Measures adopted this value to define c_0; as a result, several stabilized lasers can now act as length standards with much greater accuracy than the historical krypton-86 lamp. *See* LIGHT; PHYSICAL MEASUREMENT; WAVELENGTH STANDARDS.

Characteristics and performance of frequency standards

Characteristic	Ovened quartz oscillator	Rubidium gas cell	Cesium beam	Hydrogen maser
Frequency f_T, MHz	0.01–100	6834.682614	9192.631770	1420.495752
Q factor	10^6	10^7	10^7	10^9
Reproducibility	Must be set	10^{-10}	2×10^{-12}	10^{-12}
Aging rate per year	10^{-8}	$<10^{-10}$	$<10^{-13}$	10^{-12}
Stability per day	10^{-11}	3×10^{-13}	2×10^{-13}	10^{-15}
Temperature coefficient				
Device	10^{-8}	2×10^{-10}	$<10^{-13}$	10^{-11}
Instrument	10^{-10}	10^{-12}	$<10^{-13}$	10^{-15}

The frequency chain shown in Fig. 5 is now mainly of historical interest. Later work has been directed toward the extension of measurements into the visible region of the spectrum, approximately 400–700 THz. The critical requirement to select a set of laser wavelengths which are nearly harmonically related, so that the beat frequencies are within the range of microwave counter systems, has been relaxed by the development of the optical-frequency comb generator. This consists of an electrooptic crystal, coated with reflecting surfaces to produce a Fabry-Perot cavity tuned to the input laser frequency and placed in a microwave resonator. It generates a comb of several hundred sidebands around the laser frequency with separations equal to the microwave frequency. As in the case of the microwave counter, the precise harmonic which produces the required beats for the counting system may be determined by shifting the microwave frequency by a small amount and observing the change in the count. *See* CAVITY RESONATOR; ELECTROOPTICS; FREQUENCY COMB; INTERFEROMETRY.

Frequency standards. To calibrate the internal quartz oscillators in frequency counters, and to enable frequency measurements to be made at the highest accuracies, up to and occasionally beyond 10^{-12}, standards laboratories require a selection of very stable frequency standards. The four types in common use are the temperature-stabilized or ovened quartz crystal oscillator, the rubidium gas cell, the cesium atomic beam standard, and the hydrogen maser. Their performance (see **table**) depends essentially upon the quality factor Q—the ratio of the resonant or transition frequency f_T to its half-bandwidth—and the sensitivity of f_T to changes in the properties of materials or in the ambient and operating conditions. A high value of Q means that the system acts as a narrow filter, reducing the effects of random noise processes and improving the short- to medium-term stability. *See* Q (ELECTRICITY).

The long-term stability is determined by the variation of f_T with time. In only one type of standard, the quartz crystal oscillator, is f_T determined directly by bulk physical properties; as a result, this has the poorest long-term stability of the four. However, quartz oscillators are employed in most of the atomic standards to reduce the short-term noise and to provide a convenient output frequency (usually 10 MHz). In these, f_T is set by atomic transitions whose properties are in principle fixed and universal. In practice, small interactions with the containment system and the operating conditions mean that this ideal is not completely realized. In the rubidium gas cell, the transition is perturbed by collisions with other buffer gas atoms whose temperature and composition may change in time; in the hydrogen maser, collisions of the hydrogen atoms with the inert coating inside the containing bulb produce the so-called wall shift, which depends upon the condition of the coating. Atoms in the cesium beam standard are very well isolated from each other and the container, and this is reflected in the low drift rates and temperature coefficients observed. *See* ATOMIC CLOCK.

The effects of these factors are well illustrated by a comparison (**Fig. 6**) of the long-term performances against Universal Coordinated Time of a high-stability cesium standard and a good-quality hydrogen maser. The cesium frequency stability is limited by random noise processes as result of its lower Q, but shows little evidence of long-term drift. The maser, which has a very high Q, is very stable in the short term but possesses a linear frequency drift. Small adjustments to the maser operating conditions, as made around day 800, can alter the drift rate significantly.

Because drift rates can be measured and corrections made, the factors limiting the accuracy achieved with a frequency standard are the type and magnitude of the various internal noise sources. Plotted as a function of the measurement time (**Fig. 7**), the stability of a frequency standard shows a general boat-shaped curve. At short times, where white phase noise dominates, the stability improves as the measurement time increases. It eventually flattens

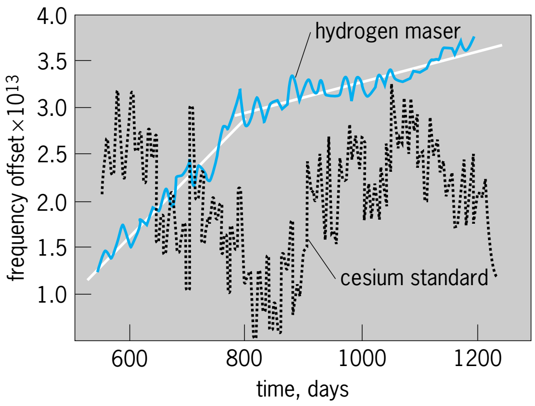


Fig. 6. Long-term stabilities of a cesium standard and a hydrogen maser. Straight lines indicate a linear drift of the maser.

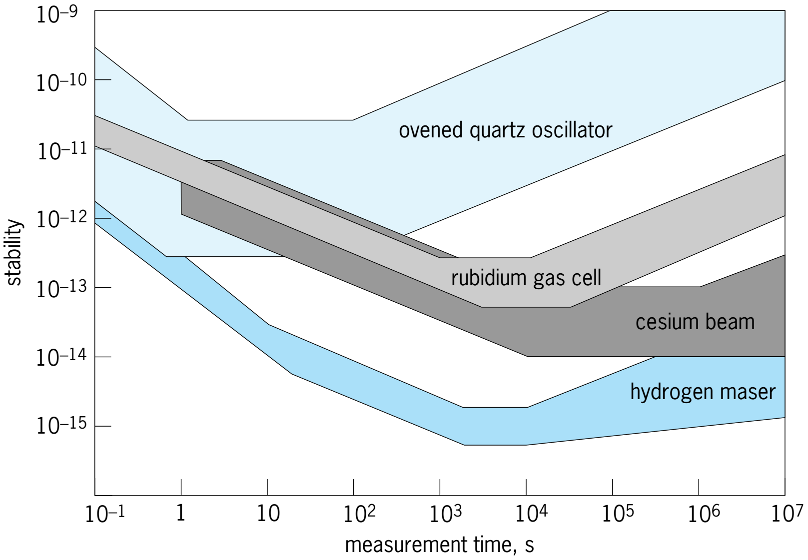


Fig. 7. Stability of frequency standards as a function of measurement time. The spread for each type indicates the range of values expected from good commercial standards. The graph is idealized to emphasize the change, with measurement time, of the type of noise that limits performance.

off to an optimum value, when flicker noise is the major component, before rising again as random-walk phase noise sources become important. *See* ELECTRICAL NOISE.

Comparison techniques. The comparison of frequency standards benefits from the fact that the signals are close to ideal: they are clean and stable, and the differences from the nominal value, usually 1, 5, or 10 MHz, are extremely small. The beat frequency is often too low to be measured directly; for example, a difference of 10^{-10} at 10 MHz corresponds to a frequency of 0.001 Hz. The resolution may be improved by measuring the phase of the beat frequency or by employing a frequency difference multiplier.

The relative phase θ of two sine waves may be measured by converting them to square waves and measuring the overlap of the waveforms with AND gates or J-K bistable logic elements. Digital methods have several advantages over analog techniques: the output varies linearly with θ and is independent of the signal amplitudes, there is no ambiguity in the sign of the frequency difference, and the constant of proportionality between output voltage and phase is easily determined. The main disadvantage of the simpler digital circuits is that there is often a dead zone around zero phase difference due to the finite transition time of the logic circuits. This may be removed by mixing a third frequency with both the unknown and reference signals. *See* LOGIC CIRCUITS.

In a frequency difference multiplier (**Fig. 8**), the unknown frequency is multiplied by a factor of 10, while the reference frequency is multiplied by 9. When the two are subtracted in a mixer, the frequency difference Δf has been increased by a factor of 10. Several stages can be placed in series to increase the multiplication factor. The output signal is measured with a conventional counter. As Δf is superimposed upon the reference frequency, there is no ambiguity as to its sign. For a four-stage multiplier and an unknown frequency of approximately 1 MHz, the resolution on the highest range is 10^{-10} for a measurement time of 1 s. It may be improved to below 10^{-13} for longer times as long as the signals are clean and stable. Ultimately the practical limit to the multiplication factor is set by noise in the multiplication and mixing circuits.

Sine-wave signals contain a large measure of redundancy; for the 10-MHz output from a standard with an accuracy of 10^{-12}, the number of complete cycles is known for periods of up to a day. Converting continuous frequency signals into timing pulses allows concentration on the small differences accumulating over long periods. An additional advantage, particularly important for signals being transmitted or broadcast over long distances in the presence of interference, is that signal continuity is no longer necessary.

Comparisons with remote standards. For measurements and calibrations at the highest level of accuracy, working standards must be referred to the national or international standards of time and frequency. The International Atomic Time (TAI) scale provides the

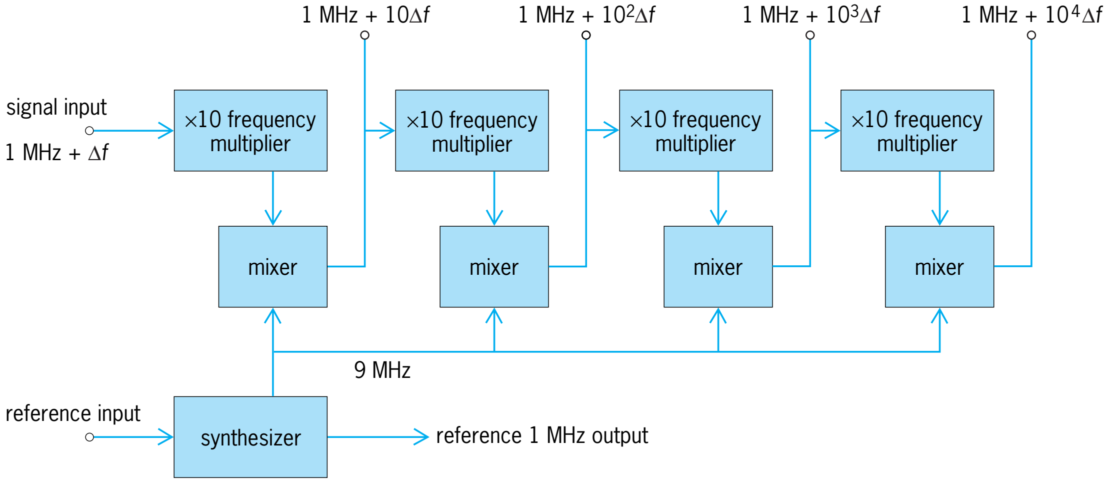


Fig. 8. Frequency difference multiplier. The unknown frequency of the signal input is approximately 1 MHz.

ultimate reference in this respect. It is generated at the International Bureau of Weights and Measures (BIPM) in Paris from the readings of a worldwide set of commercial cesium clocks, long-beam cesium primary standards, and hydrogen masers. The results are issued in the form of Universal Coordinated Time (UTC), which contains adjustments to TAI in the form of leap seconds added every 1–2 years, which keep it close to the astronomical time scale based upon the rate of rotation of the Earth. Because TAI and UTC are calculated "paper" scales with values for each month issued 1 month in arrears, the national standards laboratories in many countries generate real-time approximations to UTC. These are "steered" so that the difference from UTC is not more than 100 ns in time and around 10^{-13} in frequency.

At one time, laboratory standards could be compared with national or international standards through traveling cesium clocks. This system is time-consuming, is expensive, and allows only infrequent comparisons. It has therefore largely been replaced by day-to-day monitoring against a remote reference standard using suitable transmissions, for example, from very low frequency (VLF) or low frequency (LF) standard broadcasts, the Loran-C navigational chain, television signals, and the satellite Global Positioning System (GPS). Suitable transmissions are also monitored by national laboratories, which publish tables of corrections to enable users to refer their measurements to the national standards. The main uncertainty arises from variations in the propagation delays between the transmitter and different laboratories. *See* LORAN; SATELLITE NAVIGATION SYSTEMS.

Phase measurement has been applied to the monitoring of laboratory standards using the VLF or LF services provided by some national standards laboratories. A phase-tracking receiver detects, amplifies, and filters the signal. A higher frequency, usually 1 MHz, is then generated by a phase-locked loop, and the phase difference from the corresponding frequency from the laboratory standard is monitored. The phase stability of the received signal is sensitive to interference of the ground-wave component by the variable sky wave, limiting the useful range to around 600 mi (1000 km) from the transmitter. Measurements are best carried out around the middle of the day, when the sky-wave effects are at a minimum. A resolution equivalent to a mean frequency change of 10^{-12} from one day to the next can be achieved, although propagation errors in the middle of winter may be several times larger.

The transmissions from the satellites forming the GPS are particularly convenient for this purpose, as they contain timing information from the atomic standard present in each satellite, and the signal may be received at any point on the Earth's surface. The pseudo-random-noise-coded signals may be analyzed by a suitable receiver to give time with respect to the U.S. Naval Observatory's realization of UTC. Although UTC(USNO) is very stable in the long term, frequency measurements may be affected by short-term noise. Common-view operation, in which the same schedule of GPS satellites is observed by the user and a national standards laboratory, allows cancellation of the short-term differences. In this case, frequency measurements with respect to the national standard accurate to 10^{-13} over one day may be achieved. The comparison process has been simplified by the development of GPS-disciplined oscillators, that is, instruments in which a quartz or rubidium standard is controlled by reference to the signals from GPS satellites to improve the long-term performance.

Peter B. Coates

Bibliography. A. E. Bailey (ed.), *Microwave Measurements*, 2d ed., 1989; K. M. Baird, Frequency measurement of optical radiation, *Phys. Today*, 36(1):52–57, January 1983; Chronos Group, *Frequency Measurement and Control*, Micowave Technology 7, 1994; V. E. Bottom, *Introduction to Quartz Crystal Units Design*, 1982; J. Jesperson and J. Fitz-Randolph, *From Sundials to Atomic Clocks: Understanding Time and Frequency*, 2d ed., 1999; P. Kartaschoft, *Frequency and Time*, 1978; M. Ohtsu (ed.), *Frequency Control of Semiconductor Lasers*, 1996; J. Vanier and C. Audoin, *The Quantum Physics of Atomic Frequency Standards*, 2 vols., 1989.

Frequency modulation

A technique used in telecommunications transmission systems whereby the instantaneous frequency of a carrier signal is determined by the characteristics of an information signal, the so-called modulating signal. Frequency modulation (FM) is a form of angle modulation. For systems in which the modulating signal is digital in nature, the term frequency-shift keying (FSK) is usually employed. *See* ANGLE MODULATION.

The mathematical properties of FM signals were first presented by J. R. Carson in 1922. Carson established the fact that frequency-modulated carrier signals generally occupy a larger transmission bandwidth than would be the case for amplitude-modulated carriers. Further investigation by E. H. Armstrong demonstrated the improved noise suppression properties of FM as a benefit of the expanded bandwidth. Thus, by trading increased transmission bandwidth for improved signal-to-noise ratio it is possible to reduce transmitter power requirements for a given transmission range or desired performance. This property is of special importance in applications such as satellite communications, where satellite transmitter power and weight are important considerations. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); COMMUNICATIONS SATELLITE; SIGNAL-TO-NOISE RATIO.

Additional advantages of FM over other forms of modulation are due to the constant amplitude of the modulated carrier, independent of the modulating signal. For example, this feature provides partial immunity to the presence of nonlinearities in the signal processing equipment and allows the reduction of amplitude variations due to signal fading and additive noise in the transmission channel.

Constant-amplitude signals also exhibit constant levels of peak and average power. This property simplifies the design and implementation of transmitters, amplifiers, and automatic volume control circuits, and mitigates against problems associated with circuit nonlinearities.

Carrier signal specification. In the usual definition of FM, the carrier signal is taken to be a periodic function of time represented mathematically as Eq. (1).

$$c(t) = A\cos(2\pi f_c t + \theta) \qquad (1)$$

Here the quantities A, f_c, and θ are constants known as the amplitude of the carrier signal in volts, its frequency in hertz, and its phase angle in radians, respectively. *See* CARRIER.

If θ is allowed to be a function of time $\theta(t)$, $c(t)$ will not be periodic and the concept of frequency would, strictly speaking, not apply. In this case it is meaningful to define the concept of instantaneous frequency $f_i(t)$, given by Eq. (2).

$$f_i(t) = f_c + \left(\frac{1}{2\pi}\right) d\theta(t)/dt \qquad (2)$$

Clearly this quantity varies with the time rate of change of the carrier signal's phase. When time is measured in seconds, instantaneous frequency is again measured in hertz.

The difference between $f_i(t)$ and f_c is known as the frequency deviation Δf and is given by Eq. (3).

$$\Delta f = f_i(t) - f_c = \left(\frac{1}{2\pi}\right) d\theta(t)/dt \qquad (3)$$

Carrier modulation. Let $m(t)$ be the information signal modulating the carrier and assume its maximum value to be m_p and its bandwidth to be B hertz. In frequency modulation $m(t)$, which is assumed to be varying continuously with time, is impressed upon the instantaneous frequency of the carrier signal according to Eq. (4), so that the frequency deviation is given by Eq. (5).

$$f_i(t) = f_c + km(t) \qquad (4)$$

$$\Delta f = km(t) \qquad (5)$$

Here k is the so-called deviation constant. The maximum frequency deviation is then given by Eq. (6).

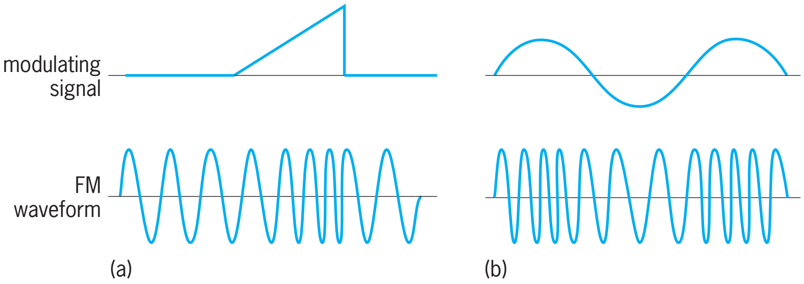


Fig. 1. Examples of modulating signals and corresponding frequency-modulated waveforms. (*a*) Ramp modulating signal. (*b*) Periodic modulating signal.

$$\Delta f_{\max} = km_p \qquad (6)$$

Combining Eqs. (1), (2), and (4) shows that the modulated carrier takes the form given by Eq. (7).

$$s(t) = A\cos\left[2\pi f_c t + 2\pi k \int m(\tau)d(\tau)\right] \qquad (7)$$

For the case of frequency-shift keying, $m(t)$ is a discrete waveform that can assume only a finite number of amplitudes that can change only at discrete points in time. These amplitudes are impressed onto the carrier signal by shifting its frequency by predetermined amounts. In binary FSK the modulated carrier signal corresponding to the two amplitudes takes the form given by Eqs. (8). Here Δf is known as the frequency deviation.

$$s_1(t) = A\cos[2\pi(f_c + \Delta f)t] \qquad (8a)$$

$$s_2(t) = A\cos[2\pi(f_c - \Delta f)t] \qquad (8b)$$

Figure 1 shows examples of frequency-modulated periodic carrier signals for which the modulating signal $m(t)$ is either a ramp function (Fig. 1*a*) or is itself a periodic function of time (Fig. 1*b*).

For a sinusoidal modulating signal of the form $m(t) = a\cos(2\pi f_m t)$, where a is a constant, the modulating index β is defined by Eq. (9), and represents the ratio of maximum frequency deviation relative to the frequency of the modulating signal. For binary FSK the modulation index is defined by Eq. (10).

$$\beta = ka/f_m \qquad (9)$$

$$\beta = \Delta f/B \qquad (10)$$

Spectrum and power characteristics. Theoretically, the energy spectrum of a frequency-modulated signal exhibits an infinite number of sidebands and so extends over the entire frequency spectrum. Practically, however, only a finite number of sidebands contain significant energy. Although the computation of the effective bandwidth is quite complicated, certain approximations can be made that lead to simple formulas. Two cases are typically distinguished. In so-called narrowband FM, the maximum frequency deviation is assumed to be quite small, in which case it can be shown that the effective bandwidth of the FM signal is given by Eq. (11).

$$\mathrm{BW}_{\mathrm{NBFM}} = 2B \qquad (11)$$

See SIDEBAND.

The case where the above assumption does not hold is referred to as wideband FM. For large values of frequency deviation the bandwidth is essentially determined by the maximum frequency deviation, and is given by Eq. (12).

$$\mathrm{BW}_{\mathrm{WBFM}} = 2\Delta f_{\max} \qquad (12)$$

These two special cases can be combined into the general formula (13) which is known as Carson's

$$\mathrm{BW}_{\mathrm{FM}} = 2(\Delta f_{\max} + B) \qquad (13)$$

rule. A similar equation holds for the bandwidth of a binary FSK-modulated carrier.

The average power of the modulated signal, as well as its peak power, are constant, and relative to a 1-ohm resistance they are equal to $A^2/2$ watts and A^2 watts, respectively. Both powers are apportioned between the unmodulated carrier and the modulating signal. Thus, as the power in the modulating signal increases, the power in the unmodulated carrier decreases correspondingly.

Modulation devices. There are several techniques for generating a frequency-modulated signal.

The direct method of modulation typically employs a voltage-controlled oscillator whose instantaneous frequency can be changed linearly in accordance with a voltage applied at its input. Such devices usually include a resonant *LC* circuit with either a varactor diode or a reactance tube whose capacitance or inductance can be changed in direct proportion to an applied voltage, thereby changing the resonant frequency of the circuit (**Fig. 2**). In some cases a certain amount of frequency multiplication may be necessary to produce the required amount of frequency deviation. *See* FREQUENCY MULTIPLIER; RESONANCE (ALTERNATING-CURRENT CIRCUITS).

In the indirect method of generating a frequency-modulated signal, a two-step process is followed. First, the modulating signal $m(t)$ is passed through an integrator whose output is multiplied with the unmodulated carrier signal.

Adding to that product the carrier signal phase-shifted by 90° results in the so-called narrowband FM signal. In the second step, the frequency of the narrowband signal is increased through one or more steps of frequency multiplication and spectrum translation (**Fig. 3**). *See* FREQUENCY MODULATOR.

Demodulation. As shown by Eq. (4), demodulation of the frequency modulated signal $s(t)$ in Eq. (7) requires the determination of its instantaneous frequency $f_i(t)$. One commonly used method of extracting $f_i(t)$ involves the intermediate generation of an amplitude-modulated signal whose amplitude is proportional to $f_i(t)$, and the subsequent use of an AM demodulator to recover the original modulating signal $m(t)$.

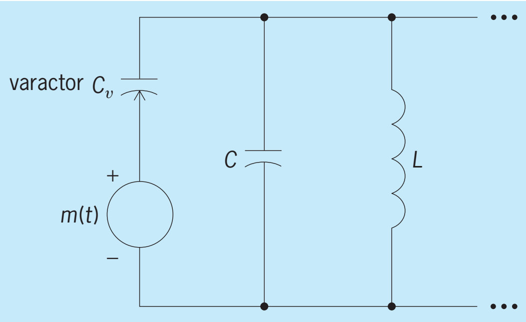


Fig. 2. Example of direct method of frequency modulation.

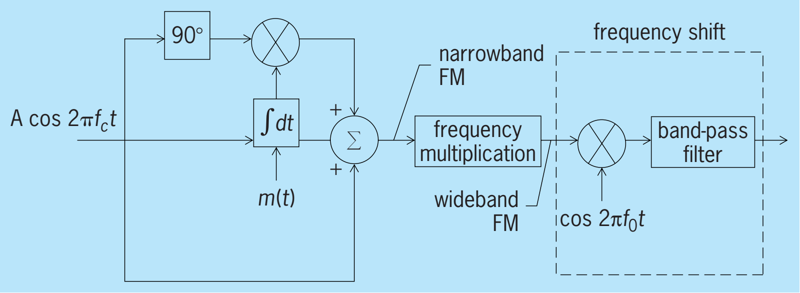


Fig. 3. Example of indirect method of frequency modulation.

The FM-to-AM conversion is generally performed by devices known as discriminators, of which several types are in common use. One of them performs the differentiation of $s(t)$ with respect to t, which results in Eq. (14).

$$ds(t)/dt = -2\pi A[f_c + km(t)] \times \sin\left[2\pi f_c t + 2\pi k \int m(\tau)d(\tau)\right] \qquad (14)$$

This quantity represents a frequency-modulated carrier whose amplitude varies in proportion to $f_i(t)$. In another method $s(t)$ is processed through a filter whose gain characteristic is linear over the range of frequencies covered by the signal. Such linear gain characteristics can be achieved in circuits known as balanced discriminators. *See* ELECTRIC FILTER.

Standard amplitude demodulation techniques such as envelope detection can be used on the output of the discriminator to recover $m(t)$. *See* AMPLITUDE-MODULATION DETECTOR.

As an alternative to FM-to-AM conversion, FM demodulation can be achieved through the use of phase-locked loops or other so-called feedback demodulators. These generally are of greater complexity, but perform better in low signal-to-noise ratio environments. In such feedback systems, the difference between the output of the demodulator and a reference signal is applied at the input, which in turn increases or decreases the output, depending on whether the feedback is positive or negative. The result is generally a decrease of distortion, especially for narrowband signals exposed to wideband channel noise or fading. *See* PHASE-LOCKED LOOPS.

Instantaneous frequency can also be determined through the technique of measuring the rate at which the amplitude of the modulated carrier changes its polarity. Devices operating on this basis are known as zero-crossing detectors. *See* FREQUENCY-MODULATION DETECTOR.

Performance. Since the information in a frequency modulation system is contained in the frequency of the transmitted signal, distortions of the signal during transmission that affect its amplitude can be removed prior to demodulation through a process known as amplitude clipping. This technique is particularly effective in the presence of additive noise. Since clipping tends to produce higher-order harmonic frequencies, low-pass filtering after demodulation is frequently employed.

Wideband FM provides a tradeoff between the bandwidth occupied by the transmitted signal and the effective postdemodulation signal-to-noise ratio. As an example, it can be shown that for additive channel noise with a constant power spectral density and a large ratio of carrier power to modulating signal power, the signal-to-noise ratio increases as the square of the transmission bandwidth. This effect, however, no longer holds when the signal-to-noise ratio falls below a certain limit. In that case, the receiver locks unto the noise as the primary signal and suppresses the information signal. This phenomenon is known as noise capture and is sometimes referred to as the threshold effect. The signal-to-noise ratio at which it occurs is known as the threshold. Under certain assumptions it can be shown that the onset of noise capture occurs at a signal-to-noise power ratio of 10.

Preemphasis and deemphasis. An additional improvement in the signal-to-noise ratio beyond the expansion of the signal bandwidth can be achieved by taking advantage of the fact that the noise power at the receiver output increases parabolically with increasing frequency, even if the receiver input noise is spectrally flat. The high-frequency components of the information signal are therefore more affected by the noise than the low-frequency components. By artificially boosting the high-frequency components of the information signal prior to modulation and restoring them to their original value after demodulation, a signal-to-noise ratio advantage can be gained. Circuits that perform these operations are known as preemphasis and deemphasis filters.

Applications. Due to its enhanced performance over amplitude modulation, wideband FM is widely used in commercial radio broadcasting, although at the cost of an increase in occupied bandwidth to 200 kilohertz, compared to 10 kHz for commercial AM broadcasting. The additional bandwidth allows for the transmission of audio program material over a range of 15–20 kHz compared to only 5 kHz for AM. It is also possible to transmit audio signals in stereo through the multiplexing of two audio channels, and to include additional channels of monophonic or stereo programming material for special purposes, all within the same 200-kHz bandwidth. *See* FREQUENCY-MODULATION RADIO; STEREOPHONIC RADIO TRANSMISSION.

Wideband FM is also used in the transmission of the audio portion of commercial television programs, and finds widespread application in satellite communications for telephony, single-channel-per-carrier systems, and wide-area commercial television distribution. *See* TELEVISION.

Narrowband FM is primarily employed in telemetry and mobile communication systems. Its major advantage over AM lies in its better response at very low frequencies and its ability to reject large additive noise pulses through the technique of amplitude limiting. *See* MOBILE RADIO; MODULATION; TELEMETERING.

Hermann J. Helgert

Bibliography. B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3d ed., Holt, Rinehart and Winston, 1998; J. Proakis and M. Salehi, *Fundamentals of Communication Systems*, Pearson Prentice Hall, 2005; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, 1990.

Frequency-modulation detector

A device for the detection or demodulation of a frequency-modulated (FM) wave. Frequency-modulation detectors operate in several ways. In one class of detector, known as a discriminator, the frequency modulation is first converted to amplitude modulation, which is then detected by an amplitude-modulation detector. Another type of FM detector employs a phased-locked oscillator to recover the modulation. A still different type converts the frequency modulation to pulse-rate modulation, which can be converted to the desired signal by use of an integrating circuit. Examples of these types will be discussed here. *See* AMPLITUDE-MODULATION DETECTOR.

Discriminator. An amplitude-modulation (AM) detector demodulates an FM wave, provided the detector is detuned so that the carrier frequency is on one side of the passband instead of in the center of the passband. Frequency variations are then converted into amplitude variations or modulation (**Fig. 1**), and an AM detector produces the desired output signal. This detection technique has serious limitations, however. Low distortion is realized if the amplitude variation is linearly related to the frequency variation or deviation of the FM wave. This requirement confines the frequency deviation to a linear or straight-line portion of the tuned-circuit response curve. This restriction seriously limits the permissible deviation, and thus nullifies the advantages of FM that result from high deviation ratios. However, amateur radio operators sometimes use this detuning technique to receive narrow-band FM on an AM receiver. *See* FREQUENCY MODULATION.

The technique of converting the frequency variations of an FM signal into amplitude variations is used in the Foster-Seely discriminator and the ratio

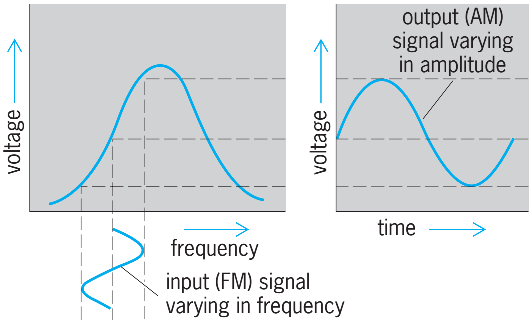


Fig. 1. Conversion of frequency modulation to amplitude modulation by a detuned resonant circuit.

detector. In these circuits the primary voltage of the transformer used to couple the final intermediate-frequency (i-f) amplifier or limiter is coupled to the center tap of the tuned secondary, which is connected to two detector diodes in series. These discriminators performed well but were not easily adaptable to integrated circuits and have, therefore, been essentially eliminated from modern radio technology.

Quadrature detectors. Frequency-modulated waves may be demodulated by a phase-shifting network, an analog multiplier, and a low-pass filter (**Fig. 2**). The *RLC* circuit in the phase shifter is tuned to the carrier frequency of the FM wave, and the reactance of the series capacitor C_1 is large in comparison with the resistive impedance of the tuned circuit at the carrier frequency. Then the voltage v_2 is in phase with the capacitive current, which is approximately 90° out of phase with the input voltage v_i at the carrier frequency. Therefore, if v_i is a sine wave, v_2 is a cosine wave. When the input signal is v_i is modulated, the input frequency changes by $\delta\omega$; thus the instantaneous phase is $(\omega_O + \delta\omega)t$. At this higher frequency the tuned circuit is no longer resonant but is capacitive, because the capacitive current exceeds the inductive current and the voltage v_2 lags the capacitive current and its phase leads the phase of v_i by less than 90°. If the relative phase shift is proportional to the frequency deviation $\delta\omega$, then the result of multiplying the two points is prefilter output v'_0, which is the sum of a low-frequency and a high-frequency term. The high-frequency term is eliminated by the low-pass filter, leaving an output voltage v_0 which differs from the modulating voltage only in amplitude. *See* ELECTRIC FILTER.

The assumption that the relative phase shift of the phase-shifting network is directly proportional to the frequency deviation of the modulated wave is not strictly true. However, the distortion resulting from this lack of linearity may be limited to about 1% if the bandwidth of the resonant circuit in the phase-shifting network is at least five times as wide as the maximum deviation of the FM wave.

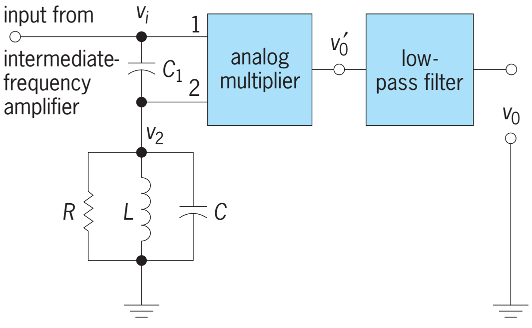


Fig. 2. Quadrature detector.

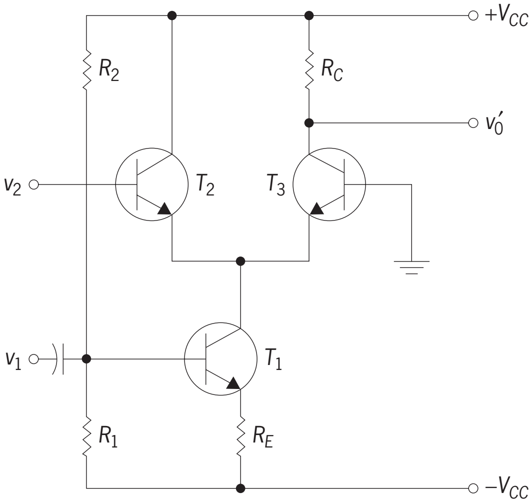


Fig. 3. Elementary multiplier circuit.

An elementary multiplying circuit is given in **Fig. 3**. This circuit is essentially the same as the unbalanced differential, or emitter-coupled, amplifier except that the base of the current-source transistor T_1 is capacitively coupled to an input terminal. Therefore, the collector current of T_1 is controlled by the input voltage v_1. The signal component of this current is essentially v_1/R_E. On the average, one-half of this current flows through T_2 and the other half through T_3. The collector current of T_3 due to a signal v_1 is therefore $v_1/2R_E$, and the voltage gain of a common-base or common-emitter transistor amplifier is g_mR_L, where g_m is the transconductance and R_L is the load resistance of the capacitor, here equal to R_C. For a transistor at normal temperature, it may be shown that $g_m = 40I_C$, leading to Eqs. (1) and (2).

$$\frac{v'_0}{v_2} = g_mR_C = 40I_CR_C = \frac{40v_1R_C}{2R_E} \qquad (1)$$

$$v'_0 = \frac{20R_C}{R_Ev_1v_2} \qquad (2)$$

See AMPLIFIER; DIFFERENTIAL AMPLIFIER.

Integrated four-quadrant multipliers based on the above principle are used extensively in various types of demodulators. These multipliers accept signal voltages of either polarity at both input ports.

The analog multiplier may be replaced by an AND gate in the quadrature detector. The gate is then called a coincident circuit. When the carrier is unmodulated, the 90° phase shift causes the two gate inputs to be positive and the gate output to be high for about 90° out of each cycle with average output voltage of approximately $V_{CC}/4$, where V_{CC} is the collector supply voltage. When the modulation causes the frequency to increase, the relative phase decreases and the two inputs are high, or coincident, for more than 90° out of each cycle; the average output increases in proportion to the frequency deviation toward a limiting voltage of $V_{CC}/2$. On the other hand, as the frequency deviation is downward, the coincidence time decreases toward zero as the relative phase approaches 180° and the output voltage decreases toward zero. The relative phase shifts must, however, be small in comparison with these limiting values in order to achieve acceptable distortion levels. The no-modulation phase shift in the phase-shifting network may differ considerably from the ideal 90° without serious degradation of the detector output. *See* LOGIC CIRCUITS; SWITCHING CIRCUIT.

Locked-oscillator detector. Another type of FM detector is the locked oscillator, in which a local oscillator follows, or is locked to, the input-signal frequency. The phase difference between the local oscillator and the input signal is proportional to the frequency deviation of the input signal. A phase detector and low-pass filter are used to produce an output voltage that is proportional to the phase difference and therefore the frequency deviation of the input signal. The output voltage is also used to control the local-oscillator frequency and maintain its lock to the input signal over the desired frequency range. This arrangement is commonly known as a phase-locked loop (PLL).

A block diagram of a phase-locked loop is given in **Fig. 4**. The phase detector may be a multiplier or AND gate, as described above. The voltage-controlled oscillator (VCO) is basically a frequency modulator. The voltage-controlled oscillator frequency is adjusted to the carrier frequency of the desired signal. As this frequency approaches this desired signal to the point where the difference frequency is within the bandwidth of the low-pass filter, the voltage-controlled oscillator pulls into lock and is said to be captured. It maintains lock with the input signal as long as their phase difference does not exceed 90°, which is set by the limits of the phase detector. This lock range is proportional to the loop gain of the phase-locked loop. This gain, and therefore the lock range, may be controlled, or set, by inserting an amplifier between the loop filter and the output or between the output and the voltage-controlled oscillator. The choice depends upon the desired output voltage. The lock range must be greater than the maximum frequency deviation of the FM signal but should not exceed the total bandwidth allotted to the FM station. *See* FREQUENCY MODULATOR.

The phase-locked loop is a widely used demodulator for FM signals. It requires no *LC* (inductance-capacitance) tuned circuits, which are cumbersome to incorporate into integrated circuits. Furthermore, the only source of distortion in a properly locked loop is the lack of linearity of the voltage-controlled oscillator, and integrated voltage-controlled oscillators that provide distortions well under 1% are readily available. However, the phase-locked loop is more complex and therefore more costly than the simpler quadrature detectors. *See* PHASE-LOCKED LOOPS.

Integrating detector. A frequency-modulated wave can be converted to a pulse-rate-modulated wave from which the original modulating signal can be recovered by the use of an integrator. The FM input signal is first converted to a low intermediate frequency by mixing the input signal with a local oscillator signal in a nonlinear amplifier. *See* AMPLITUDE MODULATION.

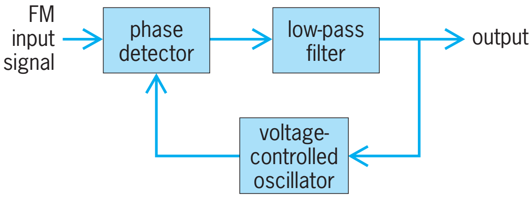


Fig. 4. Diagram of phase-locked loop.

The local oscillator frequency is nearly equal to the carrier frequency, and so the difference, or intermediate, frequency is low compared with the carrier frequency. Then the frequency deviation is a high percentage of the intermediate frequency. For example, the maximum frequency deviation of a standard broadcast FM signal at 100 MHz may be 150 kHz, or 1.5 parts per thousand. Then, if the intermediate frequency is only 200 kHz, the frequency deviation of the intermediate frequency is 750 parts per thousand. Thus, the intermediate-frequency amplifier is normally an untuned low-pass amplifier. The intermediate-frequency amplifier is followed by a limiter which clips the signal peaks and converts the sinusoidal intermediate frequency to a nearly square wave. This square wave is differentiated to obtain pulses which trigger the bistable multivibrator. The multivibrator then produces a rectangular pulse of constant amplitude and width for each input pulse. Therefore, the pulse rate is equal to the input frequency. If the frequency deviation is such that the input frequency is minimum, the pulses are widely spaced for a low-frequency input; on the other hand, the pulses are closely spaced when the deviation produces maximum input frequency. The integrator is basically a low-pass *RC* filter, which produces an output voltage proportional to the average value of the input voltage and filters out the high-frequency components of the pulses. Therefore, the output voltage is proportional to the pulse rate. The simple *RC* integrator can be replaced by an operational-amplifier-type integrator for improved filtering, linearity, and output voltage. *See* LIMITER CIRCUIT; MULTIVIBRATOR; OPERATIONAL AMPLIFIER. Charles L. Alley

Bibliography. R. G. Best, *Phase-locked Loops: Simulation, Design and Applications*, 4th ed., 1999; H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, 1980; C. A. Schuler, *Electronics: Principles and Applications*, 5th ed., 1998.

Frequency-modulation radio

Radio transmission accomplished by variation of the frequency aspect of a radio-frequency (RF) carrier signal by an input signal (called the modulating signal). The frequency-modulation (FM) technique was invented by Major Edwin A. Armstrong in the 1930s. FM radio broadcasting offers a number of advantages over amplitude-modulation (AM) radio broadcasting, most significantly higher audio fidelity and greater immunity to atmospheric and human-made sources of interference.

When an RF carrier is frequency-modulated, a theoretically infinite number of "sidebands" are created which surround the carrier (that is, are located above and below the carrier, in frequency), although in practical implementations the extent of these sidebands is limited to a finite bandwidth (called the frequency deviation). The sidebands formed during the modulation process are separated from one another by an amount proportional to the frequencies present in the audio signal. The number of significant sidebands depends on the amplitude of the modulating signal, while the amplitude of these sidebands diminishes progressively (though not monotonically) as the sidebands occur farther and farther from the carrier frequency.

Different FM systems use different frequency deviations, also called deviation ratios, depending upon the application. The greater the deviation ratio, the larger the bandwidth required for the transmission. Deviation ratios are standardized for various classes of service for optimum use of the radio-frequency spectrum by a large number of stations. For example, FM broadcast signals typically have a 75-kHz deviation frequency, resulting in a transmitted signal with a bandwidth of approximately 150 kHz (twice the deviation frequency).

Because an FM signal is constant-envelope (that is, it has a constant amplitude versus time), FM radio systems are highly resistant to the amplitude-related interference problems that plague AM radio-frequency systems. Most FM receivers are designed to remove completely any amplitude variation existing on the received signal, since these amplitude variations are by definition due to degradations and perturbations impressed upon the transmitted FM signal as it traveled from the transmit to the receive antenna. Then, the FM signal is demodulated using a frequency-discriminator detector circuit that recovers the original modulating signal, which is unaffected by the amplitude variations experienced during transmission. Another property of FM receivers is the relative freedom from interference between distant and local stations using the same channel; only the strongest signal is received, even if the wanted signal is only 3–6 dB stronger than the interfering one. This characteristic is known as the capture effect. In contrast, AM radio signals differing in strength by 35 dB result in noticeable interference. *See* AMPLITUDE MODULATION; FREQUENCY-MODULATION DETECTOR.

Since most FM receivers are equipped with a muting circuit (squelch) that silences the audio channel when no signal is present, an irritating hiss noise is not emitted from the receiver when the frequency (or channel) being monitored is not in use.

Frequency modulation is used mainly for transmissions above 25 megahertz (MHz). Typical uses are in broadcasting; television sound; mobile radio telephony; radio paging systems; space telemetry; intercity microwave relaying of all classes of public traffic including voice channels, teleprinting, facsimile, broadcast network programs, and television and computer data; and intercontinental telecommunications via satellite. Frequency modulation is used for both analog and digital communications, and phase modulation (variation of carrier phase as a function of the modulating signal) as well as frequency modulation is used (**Fig. 1**). The result, with regard to output waveform, detection, and signal-to-noise improvement, is essentially the same whether phase-modulation (PM) or FM techniques are employed. *See* PHASE MODULATION; SIGNAL-TO-NOISE RATIO.

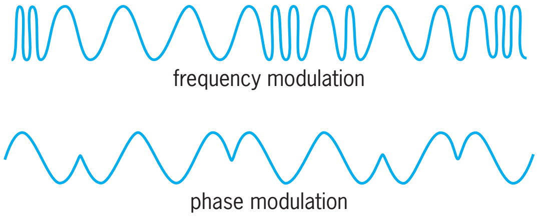


Fig. 1. Modulated waveforms.

FM broadcasting. The frequency band 88–108 MHz is allocated to FM broadcasting worldwide by international agreement. In the United States, this band is divided into 100 channels, with 200 kHz allocated for each channel. Each channel will support transmission of an analog audio signal with an audio frequency range of 50–15,000 Hz. Federal Communications Commission (FCC) rules allow for a maximum deviation of 75 kHz, and the emitted spectrum is twice this value. Because of the relatively small signal power in the modulating frequencies above 4 kHz, the received signal-to-noise ratio is improved substantially by preemphasis of the audio signal in transmission, necessitating complementary deemphasis in the receiver to restore natural program balance. In fact, preemphasis produces a sort of hybrid form of modulation, being pure frequency modulation at the lowest audio frequencies, and gradually changing to phase modulation at the highest. *See* FREQUENCY MODULATION.

Stereophonic broadcasting. Most FM stations broadcast two-channel stereophonic sound (that is, auditory perspective), where two microphones (left and right) are used in the recording of the audio source material. Stereophonic transmission must also provide normal reception for nonstereophonic (monaural) FM receivers. To accomplish this, the main channel modulation consists of combining, by additive polarities, the sum of the outputs from the left and right microphones. For stereophonic reception, a multiplexed channel is created which contains the L–R signal, amplitude-modulated onto a 38-kHz carrier signal. A 19-kHz pilot signal (38 kHz divided by 2) is sent with these L+R and L–R signals to allow for demodulation of the L–R signal (**Fig. 2**). Because the L+R and L–R signals are antiphased, high-volume peaks occur in one channel when the other is at its lowest. Each channel has access to almost half the total frequency deviation when its signal peaks to its highest volume. The available power is shared between the two, and the nonstereophonic receiver is unaffected by the multiplexed (L–R) channel and pilot.

FM subcarriers. The large channel bandwidth of FM broadcast stations permits simultaneous transmission of other services (in addition to the main-channel audio service) by use of so-called FM subcarriers. These are additional signals that are

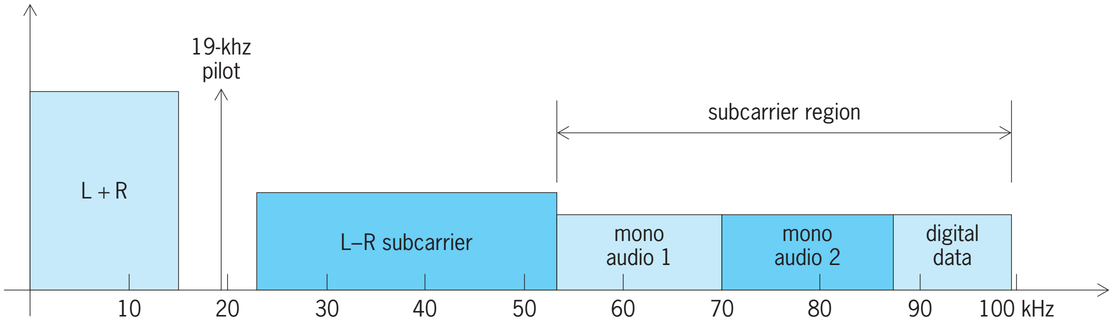


Fig. 2. FM subcarrier services.

frequency-division-multiplexed along with the main-channel audio, modulating at frequencies above the highest program frequencies. This is accomplished by frequency-division multiplexing; an example is shown in Fig. 2. FCC regulations require that the modulating signal input to an FM modulator (also called an FM exciter) be limited to the frequency band of 0–99 kHz, referred to as the FM baseband. A typical FM-baseband input is shown in Fig. 2, consisting of a main-channel audio signal (stereo, consisting of L+R, 19-kHz pilot, and L–R subcarrier) and, in this example three additional FM subcarriers, occupying frequencies from approximately 53 kHz to 99 kHz in the FM baseband. These subcarriers can be used for subscription or other special services. Normal FM broadcast receivers will not respond to these subcarrier signals, but special receivers are employed to select at least one of them, to the exclusion of the normal broadcast program. These multiplexed special emissions are permitted by the FCC under special communication authorization (SCA). *See* RADIO BROADCASTING; STEREOPHONIC RADIO TRANSMISSION.

In-band/on-channel (IBOC) digital radio. FM radio broadcasters in the United States were granted permission by the FCC in October 2002 to begin transmitting a digital radio signal along with their analog FM signal using a technique known as hybrid in-band/on-channel (IBOC) digital radio broadcasting (**Fig. 3**).

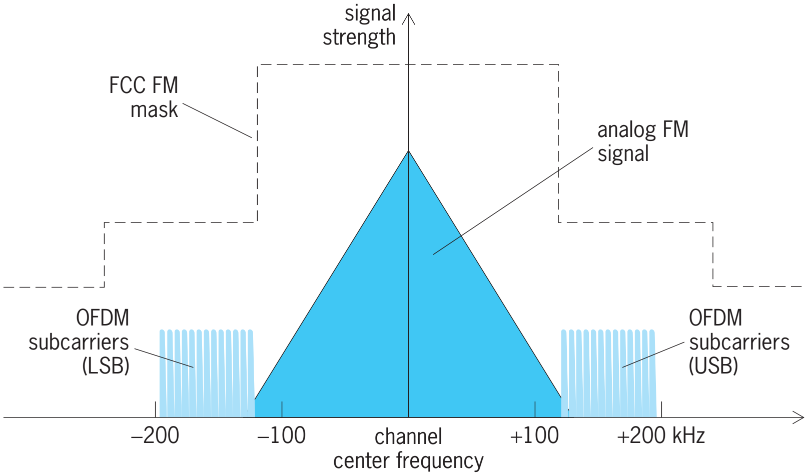


Fig. 3. IBOC digital radio signal (FM hybrid).

In an FM IBOC signal, the main-channel audio signal is broadcast (simulcast) in both the analog and digital portions, and the receiver uses the analog signal as a backup for the digital signal, which is less powerful and cannot be received as far away from the transmitter as can the analog FM signal. The digital portion of the IBOC signal uses a digital modulation technique called orthogonal frequency-division multiplexing (OFDM), which is a very robust technique favored for use by terrestrial digital broadcasting systems.

FM mobile transmission. Millions of land, maritime, aeronautical mobile, and portable FM transceivers are used by police, fire fighters, public safety agencies, industrial and commercial enterprises, private citizens, and radio amateurs who desire the benefits of en route telephony. Prior to the widespread use of commercial cellular and personal communications services, the intensity of private two-way radio use was growing exponentially, mainly because of the availability of reliable, small, low-power-consumption solid-state equipments that are economical, and also because of the public realization of the benefits of having such communications. The rapid increase in such usage became a major concern at the time to managers of the radio-frequency spectrum when it became apparent that all available channels would be assigned in major city areas.

The channel width assigned for mobile radio telephony has been reduced in some bands to 15 kHz (12.5 kHz in the United Kingdom), from the wide-band frequency modulation initially employed, and state-of-the-art receivers provide intermediate-frequency (IF) selectivities of −100 dB for ±15-kHz channel spacings. Additional bands have been allocated for mobile radio (800 MHz), and new techniques are being investigated as an alternative to FM, such as single-sideband (SSB), spread-spectrum, stored speech, and the virtual elimination of speech by the use of data transmission in those applications where standard forms of message predominate. SSB transmission, employing linear compression and expansion to combat the adverse effects of fading on SSB systems, has demonstrated that a minimum channel separation of about 6 kHz is possible compared with 15 kHz for FM transmission. *See* SINGLE SIDEBAND.

Typically, mobile telephony is intermittent, each station transmitting only to send. Receivers are open to receive incoming calls on their channel, or the receiver is activated only when certain signaling (touch or subaudible) tones are received. Simplex transmission is the rule; that is, by means of a push-to-talk button on the handset or microphone, the transmitter is switched on to talk. Many stations can share a channel, each being informed of when the channel is clear.

Mobile transmitters have output powers of 10–50 W. Hand-held units employing powers of 1–6 W are also used. Base stations employ powers up to 250 W. Vertical polarization is used for FM radio telephony, since a mobile whip, which uses the vehicle as a ground plane, is the most practical antenna configuration. The geographical range of such stations is a function of the transmitting and receiving antenna heights. The communications range of mobile units with antenna heights of 6 ft (1.8 m) is on the order of 6 mi (3.7 km), with the range increasing to 17 mi (27 km) for mobile communications with a base station with an antenna 100 ft (30.5 m) above ground level. Favorable (flat) terrain will tend to increase communications range, while hilly or mountainous terrain and heavily treed areas (for example, tropical forests) will tend to decrease range. The use of repeaters favorably located on the top of high-rise buildings, on the top of hills, mountains, or tall towers increases the useful range for mobile-to-repeater communications to distances of 60 mi (100 km) and beyond. In the ideal situation, therefore, mobile-to-mobile communication by way of a repeater station to twice this distance is possible. The repeater receives the input signal on one frequency and rebroadcasts it on another frequency, with higher transmitter power or high antenna gain, or both. Amateur-radio 2M repeaters employ a frequency spacing of 600 kHz between input and output frequencies, necessitating high-Q cavity duplexers so that the receiver and transmitter can employ a common high-gain antenna. The transmission signal must be attenuated by more than 100 dB, even with adequate receiver IF selectivity, since the receiver must respond to signals as small as 0.2 microvolt or so in the presence of a transmission signal of tens of volts.

With increasing occupancy of the mobile radio bands by many nearby users operating on equally spaced channels, problems of intermodulation are a major limitation to further expansion in some congested areas. Intermodulation occurs when two or more frequencies combine external to, but usually in, the receiver front end to produce a difference frequency that is the same as the desired frequency. Typically, communication receivers experience intermodulation problems when the combination frequencies are about 80 dB above the receiver threshold, although improved performance is possible with state-of-the-art design.

Radio repeaters have become very sophisticated, with tone burst, subaudible tones which can be used for remote control of the repeater functions, linking, and so on, and push-button autopatch (automatic connection and push-button dialing into the public switched telephone system). The numbers of amateur-radio repeaters (simple or complex) have grown at a rate that exceeds development of commercial repeaters, and there are one or more repeaters serving most cities and surrounding areas in the United States and Canada. *See* MOBILE RADIO.

Cellular systems. Cellular systems (both analog and digital) employ equipment operating in the 800-MHz and 2000-MHz bands. The mobile user can move from cell to cell, and these movements are kept track of, whether the user is monitoring, initiating, or receiving a radio-to-telephone interconnect call. Such systems must be oriented toward microprocessor control if all the functions proposed are to be implemented. Cellular and trunking radio systems have a great deal in common. Analog modulation schemes like AMPS (advanced mobile phone system) and TACS (total-access communications system) use frequency modulation, while digital schemes like TDMA (time-division multiple access) and CDMA (code-division multiple access) use variants of phase modulation.

Radio paging service. These paging transmitters, which employ high power, a 600-W centrally located transmitter, or several 60-W transmitters frequency-synchronized to within a few hertz throughout the service area, so that reception is possible employing shirt-pocket-sized receivers and inefficient antennas inside buildings and automobiles. Tone, numeric, and text pagers are used. The sender activates the pager by telephone, modem link, or the Internet. Tone pagers signal the user to take a prearranged action. In numeric and text paging systems, the sender transmits a callback telephone number or a short text message to the user. The portable receiver or pager, selectively called by a digital or dual-tone-multifrequency (DTMF) code sequence, receives the message. The numeric or text message is displayed on a liquid crystal display (LCD). *See* ELECTRONIC DISPLAY; RADIO PAGING SYSTEMS.

Radio relaying. Frequency modulation is used for long-haul microwave radio relaying over land, over water, and to great distances using satellites, sometimes carrying thousands of simultaneous telephone conversations or several television channels. The advent of requirements for short-haul services, local distribution networks within cities, television relay, and a wide variety of optical communication services—including high-speed computer communications, electronic mail, data transmission—and other services, where it may be cost-effective to avoid the local telephone loop, presents another application for FM radio relaying.

Frequency modulation is of special importance in this application because changes to the signal amplitude along the propagation path, due to fading, do not affect reliability as significantly as with some other modulation schemes. Although the background noise increases as the transmissions fade, the signal level at the output of the receiver limiter remains steady, a highly essential characteristic in a communication system. High-efficiency amplifiers can be used on FM signals since linearity is not critical as with amplitude modulation (AM) and digital modulation.

In microwave communications, frequencies above 1 GHz and to about 40 GHz are used for terrestrial and satellite relaying. By these means, hundreds of telephone, teleprinter, television, and computer data channels of communications can be transmitted on a single microwave carrier. In large systems, many carriers operating in parallel over the same route are used to achieve increased capacity.

High communications capacities are obtained by multiplexing, using large bandwidths. Two types of frequently employed carrier modulation are analog frequency-division multiplexing-frequency modulation (FDM/FM), which is more common, and digital phase-shift keying (PSK), which is growing in use. In digital communications the analog signal is sampled at some predetermined rate, and the amplitude is quantized at this sample interval by an analog-to-digital converter. The main advantage of digital transmission consists in the possibility of overcoming noise and interference that are caused by human activities by means of special coded signals and by reception techniques which are optimum for the given conditions. In uncoded systems the signal-to-noise ratio increases linearly with increased signal bandwidth, whereas in coded systems the signal-to-noise ratio rises exponentially with increased signal bandwidth. This can be explained as follows: Doubling of the spectrum width means doubling the number of pulses (or switch transitions) in the code combination, which, according to the rules that govern combinations, squares the number of signals which can be transmitted. *See* ANALOG-TO-DIGITAL CONVERTER.

When PSK or FSK (frequency-shift keying) are used with coherent reception, high frequency and phase synchronization are essential. This is a complicating requirement for this form of modulation, particularly since the PSK signal does not normally contain a component at the carrier frequency, and thus some method of synchronization is necessary.

Radio and television broadcasting using digital communications channels is possible. Although such systems are based on the principles of speech transmission, the system parameters are different, owing to the wider band of continuous signals and the necessary increase in the number of quantizing levels. Sophisticated high-speed digital communications equipments employ not just two-level PSK (0 to 180°) but quadraphase-shift-keying (QPSK) modulation (0 or 180° as well as 90 or 270°) and octal phase-shift keying (45° between states), and thus provide nearly double or quadruple the transmission capacity.

Most systems employ analog FDM/FM. For example, a normal telephone signal occupying a range of 300–3400 Hz is used to modulate a suppressed carrier, commonly between 60 and 108 kHz. By means of electronic filters, only a single sideband is selected, this being the original signal translated to a higher frequency. These subcarriers are spaced at 4 kHz, so that 12 conversations can fill the 60–108-kHz range. In turn, the 12 channels are multiplexed with other groups of 12, to form a supergroup of 60 channels, similarly transposed. Master groups of 600, 980, 1200, 1800, or 2700 channels are further extensions of the principle.

Finally, the carrier is frequency-modulated by this aggregate of individual channels to occupy a frequency band that may reach 8 MHz. The carrier peak-deviation ratio is fractional for overland relaying. Typical deviation per telephone channel is 2 kHz. Larger per-channel deviations are used for troposcatter systems and in satellite communications. The Canadian Anik satellite system has a maximum capacity of 4800 two-way voice circuits or 10 television channels. The American *Intelsat IVA* satellite was designed to have a capacity of about 10,000 two-way voice circuits or 20 television channels. *See* COMMUNICATIONS SATELLITE; TELEPHONE SERVICE.

Video recording. In VHS video recording, the black-and-white (luminance) portion of the analog video signal is frequency-modulated. The response of magnetic video tape is nonlinear over the 3-MHz VHS video bandwidth, and FM minimizes the effects of distortion introduced during the recording process.

Telegraphy. Telegraphy, including teleprinting and binary digital data transmission, is based on frequency-shift keying (FSK), a form of FM modulation. In FSK, the modulating signal shifts the carrier frequency or its phase between two limiting values, one of which represents a mark signal and the other a space signal. This frequency shift (or phase shift) is a form of FM signaling used over a wire, cable, or radio. A common approach is to use multiplexed audio tones, each deviated a few tens of cycles (or shifted in phase between two discrete levels), to carry individual messages. A normal voice channel can carry 16–24 such telegraph channels operating at standard teleprinter speeds. High-speed data transmission requires larger bandwidths and more elaborate equipment, but nevertheless is based on the same principle of frequency or phase shift. *See* TELEGRAPHY.

Facsimile. Black-and-white images (line drawings and typed copy) can be transmitted by employing the principles used in FM telegraphy; one limit frequency corresponds to black, and the other to white, on the image to be transmitted. It is necessary only that the original document be scanned (for example, by a photosensitive cell), that the scan at the recorder be synchronized with that at the transmitter, and that some means be employed—for example, using an electrosensitive or electrostatic (dry) process—to make the recording paper black when a black signal is received and leave it white for a white signal. A continuous gray scale can be transmitted and recorded if, instead of just two frequencies, a continuous frequency shift is employed between some low frequency (say, 1500 Hz) and some higher frequency (say, 2700 Hz), the exact frequency at any instant being proportional to the gray level of the image. The recorder employs a frequency- amplitude transducer, the response of which is the inverse of that at the transmitter. More complicated modulating arrangements allow color images to be scanned and transmitted to receivers capable of color reproduction. The audio-frequency tones can be sent over telephone lines; these tones can modulate an FM radio transmitter, or they can modulate an SSB transmitter. *See* FACSIMILE.

Telemetry. Prior to digital communications, frequency modulation was the preferred method for transmission of information or data from a remote or inaccessible location such as a rocket vehicle in flight. Each condition to be remotely observed would actuate one subchannel, which, when multiplexed with other channels reporting other status conditions, modulated the radio carrier by frequency modulation. Telemetry is not much different in principle from FM and PM telegraphy, except that it is a one-way transmission system, with each channel containing a transducer which converts sensed phenomena into electrical signals. The signals may be either digitized coded pulses or continuous (analog) changes. At the receiver, each channel is suitably instrumented to record directly in real time the information reported by the transducer on that channel; or the FM signal can be frequency-translated and recorded directly on magnetic tape for subsequent analysis. Frequency or phase modulation is very useful for digital data recording, since, for example, phase modulation is effective against varying oxide thicknesses on the magnetic tape and varying signal strength. Telemetry systems exist in all degrees of sophistication, up to hundreds of data channels. *See* TELEMETERING.

John S. Belrose; David H. Layer; Robert W. Denny, Jr.

Bibliography. E. Armstrong, Radio Signaling System, U.S. Patent No. 1,941,066, Dec. 26, 1933; A. B. Carlson, P. B. Crilly, and J. C. Rutledge, *Communications Systems*, 4th ed., 2001; J. G. Proakis and M. Salehi, *Communication Systems Engineering*, 2d ed., Prentice Hall, 2001; M. Schwartz, *Information Transmission, Modulation and Noise*, 4th ed., 1990; H. Taub, *Principles of Communication Systems*, 2d ed., 1986; W. Tomasi, *Electronic Communication Systems: Fundamentals Through Advanced*, 3d ed., 1997; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 9th ed., 2002.

Frequency modulator

An electronic circuit or device producing frequency modulation. This device changes the frequency of an oscillator in accordance with the amplitude of a modulating signal. If the modulation is linear, the frequency change is proportional to the amplitude of the modulating voltage.

High-frequency oscillators usually employ either *LC* (inductance-capacitance) tuned circuits or piezoelectric crystals to establish the frequency of oscillation. This frequency can be controlled by changing the effective capacitance or inductance of the tuned circuit in accordance with the modulating signal. A simple example of this type of modulator is a capacitor microphone used as a capacitive tuning element in an oscillator. The acoustic waves that strike the diaphragm of the microphone change the capacitance, which in turn changes the oscillator frequency. The acoustic pressure are usually so small that the capacitance change and the frequency change, or frequency deviation, are linearly related to the acoustic pressures. However, this modulation method is cumbersome and seldom used. Practical circuits usually employ a varactor diode to change the oscillator in accordance with a modulating voltage.

Varactor modulator. The oscillators in high-frequency electronic systems, such as frequency-modulating (FM) transmitters, usually employ piezoelectric crystals for precise control of the carrier frequency. These crystals are equivalent to a series *LC* tuned circuit with an extremely high *Q*. The crystal holder has a small capacitance which is in parallel with the crystal and therefore causes parallel resonance at a slightly higher frequency than the series resonant frequency of the crystal. The actual oscillator frequency is between these two resonant frequencies and is controllable by the parallel capacitance. *See* PIEZOELECTRICITY; Q (ELECTRICITY).

The junction capacitance of a semiconductor diode varies with the diode voltage, and a reverse-biased diode may be used to control the oscillator frequency to produce frequency modulation. Low-loss diodes designed for this service are known as varactor diodes and have trade names such as Varicaps or Epicaps. A basic varactor modulating scheme is shown in **Fig. 1**. In this circuit, the transistor that drives the varactor modulator provides reverse bias as well as the modulating voltage v_m. The radio-frequency (rf) choke provides very high impedance at the oscillator frequency to isolate the transistor amplifier output impedance from the oscillator circuit but to allow the modulating signal to pass through with negligible attenuation. Only the frequency-determining part of the oscillator is shown. The symbols C_c, L_c, and R represent the electrical equivalents of the compliance, mass, and loss,

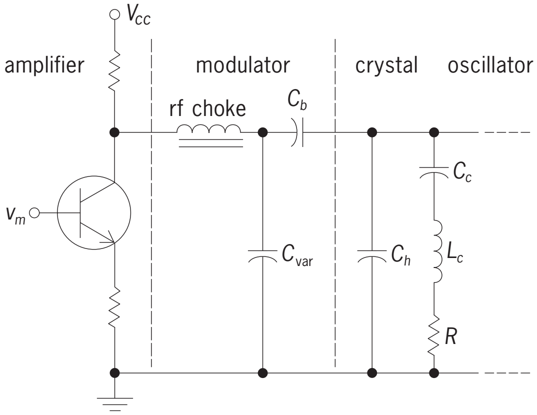


Fig. 1. Basic varactor modulator circuit. V_{CC} = collector supply voltage.

respectively, of the crystal; C_h is the crystal-holder capacitance and C_b is a dc blocking capacitor. *See* SEMICONDUCTOR DIODE; VARACTOR.

The capacitance of a typical varactor diode is inversely proportional to the square root of the reverse-bias voltage. This relationship may be expressed as in Eq. (1), where C_{var} is the varactor capacitance, $C_{-1\text{V}}$

$$C_{\text{var}} = \frac{C_{-1\text{V}}}{(-v_d^{1/2})} \tag{1}$$

the varactor diode capacitance at 1 V reverse bias, and v_d the bias voltage of the diode. Since v_d is negative for reverse bias, the quantity in parentheses is positive. This relationship is nonlinear, and even if it were linear the frequency deviation would not be linearly related to the modulating voltage, because the oscillator frequency is not linearly related to the tuning capacitance. The modulation distortion may be low for small frequency deviations. Fortunately, the distortion of the modulation system may be reduced to desirable values at normal frequency deviations by using negative feedback in the system. This feedback may be accomplished by coupling a linear FM demodulator such as a phase-locked loop (PLL) to the oscillator and feeding the voltage output back to an appropriate point in the driving amplifier where the feedback voltage has opposite polarity to the modulating-signal voltage. *See* FEEDBACK CIRCUIT; PHASE-LOCKED LOOPS.

The varactor-diode modulator is also commonly used to control the frequency of local oscillators in radio receiving equipment where programmed, push-button, or remote tuning is desirable. In these applications, conventional *LC* tuned circuits may be used.

Voltage-controlled oscillator modulator. Oscillators having a linear relationship between their output frequency and a controlling voltage have been developed for phase-locked loops. Known as voltage-controlled oscillators (VCOs), they are available in integrated circuit form as a separate chip or as a separately usable section of an integrated phase-locked loop.

The voltage-controlled oscillator is similar to an astable multivibrator. The main difference is that emitter-circuit resistances of a typical emitter-coupled multivibrator have been replaced by current sources (**Fig. 2**), which provide charge and discharge currents to the timing capacitor C_f. They are usually transistors with fixed values of collector current. In order to understand how the circuit operates, it may be assumed that transistor T_1 is cut off and I_1 is therefore the only current flowing through the timing capacitor C_f. Then the voltage across C_f changes linearly with time t at the rate given by Eq. (2). The emitter

$$v_{cf} = \frac{1}{C_f}\int_0^t I_1 dt = \frac{I_1}{C_f}t \tag{2}$$

of T_1 becomes negative with respect to the emitter voltage of T_2, which is fixed by the voltage divider R_3 and $(R_{c1} + R_4)$ between V_{CC} and the control voltage. During this time the base potential of transistor T_1 is low because transistor T_2 is either at or near saturation, the current i_3 is small, and the base potential of T_1 is near the control voltage, holding T_1 in cutoff condition. Eventually, the emitter potential of T_1 becomes lower than the base potential and T_1 begins to conduct. Therefore, the current through C_f decreases and the collector potential of T_1 drops while the collector potential of T_2 rises, thus causing the base potential of T_1 to rise while the base potential of T_2 falls toward the control voltage. This action cuts off transistor T_2 while T_1 is either at or near saturation and current I_2 flows through C_f from left to right, causing the emitter potential of T_2 to decrease linearly until T_2 begins to conduct and the switching begins. The switching action takes place very quickly and the output voltage is essentially a square wave; the voltage at each emitter is a triangular wave. As the control voltage becomes more positive, the voltage change across C_f required to cause the switching action is reduced. Therefore, the frequency of the voltage-controlled oscillator increases in direct proportion to the change in control voltage as shown in Eq. (2). *See* FREQUENCY MODULATION; MULTIVIBRATOR; OSCILLATOR. Charles L. Alley

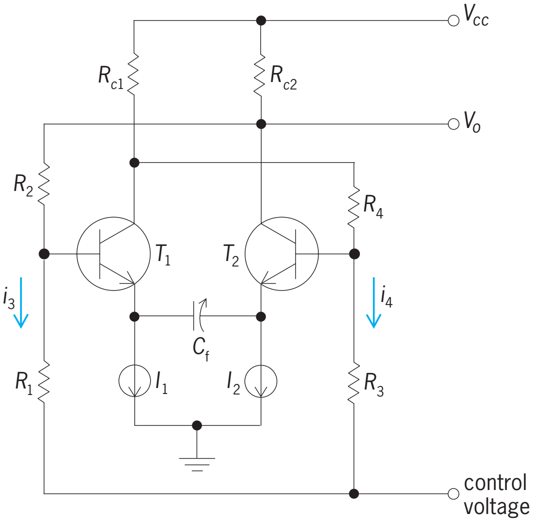


Fig. 2. Simple voltage-controlled oscillator circuit.

Bibliography. H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, 1980; A. P. Malvino, *Electronic Principles*, 6th ed., 1998; C. A. Schuler, *Electronics: Principles and Applications*, 5th ed., 1998.

Frequency multiplier

An electronic circuit that produces an output frequency which is an integral multiple of the input frequency. There are three basic types of frequency multipliers. The first type is a nonlinear amplifier that generates harmonics in its output current and a tuned load that resonates at one of these harmonics. The second type uses the nonlinear capacitance of a junction (semiconductor) diode to couple energy from the input circuit, which is tuned to the fundamental frequency, to the output circuit, which is tuned to the desired harmonic. The third type of multiplier uses a frequency divider in a locked-oscillator circuit.

Nonlinear amplifier. Harmonic-generating amplifiers are usually operated class C, biased so that the output current flows for only a small part of the input cycle, perhaps 90°. The output is then richer in harmonics, and higher efficiency is attained than with normal class C bias. The output circuit may be tuned to any harmonic, but the higher harmonics are weaker and the power output may be less than the power input if the selected harmonic is higher than the third or fourth. *See* AMPLIFIER; HARMONIC (PERIODIC PHENOMENA).

A highly efficient doubler can be devised by using two amplifiers, such as transistors, with their inputs driven with opposite polarity, obtained from opposite ends of a center-tapped coil, and their outputs connected in parallel. These amplifier-type frequency multipliers are often used in crystal-controlled high-frequency transmitters. The crystal oscillator is desirable because of its excellent frequency stability, but crystals are available for only a limited frequency range. Therefore, a moderate crystal-oscillator frequency is multiplied to a high output frequency.

Nonlinear coupler. Diode-type multipliers are usually used in very high-frequency or ultrahigh-frequency solid-state transmitters. The frequency limit at which a transistor has useful power gain is highly dependent upon the power output capability of the transistor. Varactor-diode multipliers are often used to provide an output frequency for the transmitter that is several times as high as the frequency limit of the power transistor or transistors in the transmitter. *See* TRANSISTOR.

Efficiency is of prime importance in this type of multiplier because the transmitter power output is the power output of the final transistor amplifier minus the power loss in the multiplier. Fortunately, the efficiency of varactor-diode multipliers may be as high as 80–90% if the circuit and diode *Q*s are high and the multiplication in each stage is no greater than two or three. The multipliers can be advantageously connected in cascade if higher multiplication ratios are desired. A typical varactor-diode multiplier circuit is shown in **Fig. 1**. Coil L_1 and capacitor C_1 resonate near the input frequency, while coil L_2 and capacitor C_2 resonate near the output frequency. *See* SEMICONDUCTOR; VARACTOR.

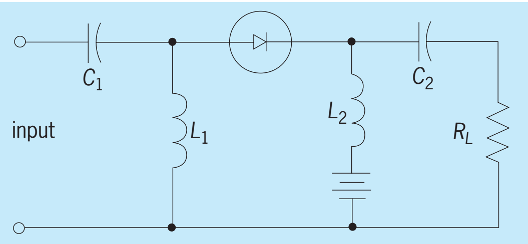


Fig. 1. Varactor-diode multiplier.

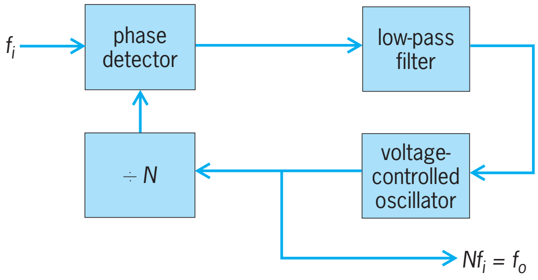


Fig. 2. Locked-oscillator frequency multiplier.

Locked-oscillator type. A convenient technique for multiplying a signal frequency by any reasonable integer *N* uses a frequency divider in a locked-oscillator circuit (**Fig. 2**). This circuit may consist simply of a phase-locked loop and a digital counter, both of which are available off-the-shelf items for output frequencies that do not exceed the capability of available high-speed counters and phase-locked loops. The input signal usually comes from a crystal-controlled oscillator because of the frequency stability of these oscillators.

Since the signal frequencies must be the same at the two phase-detector inputs (Fig. 2), the frequency of the voltage-controlled oscillator must be *N* times as high as the input-signal frequency f_i. Therefore the voltage-controlled oscillator provides the output f_o for the multiplier. This versatile type of frequency multiplier is commonly used in radio transmitters and receivers where the digital counter may be used to provide push-button selection of the communication-channel frequency. *See* DIGITAL COUNTER; FREQUENCY DIVIDER; PHASE-LOCKED LOOPS; RADIO RECEIVER; RADIO TRANSMITTER.

Charles L. Alley

Bibliography. H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, 1980; A. P. Malvino, *Electronic Principles*, 6th ed., 1998; C. A. Schuler, *Electronics: Principles and Applications*, 5th ed., 1998.

Freshwater ecosystem

A physiologically interactive system within which growth, adaptation, species composition, nutrient cycling, biological productivity, and energy flows among inland aquatic microbial, plant, and animal communities are integrated with the environment. These inland waters include lakes, reservoirs, rivers, streams, and wetlands. The discipline termed limnology evaluates how these relationships regulate, and are regulated by, natural and anthropogenic variations in the physical, chemical, and biotic environments. The standing-water (lentic) lake or reservoir ecosystem is intimately coupled with the land in its surrounding drainage area and with its running (lotic) waters that change chemical components from the land as they are transported to the lake. Often the land-water interface supports

TABLE 1. Water in the biosphere[*]

	Volume, 10^3 km ^{3}	Percent of total	Renewal time
Oceans	1,370,000	97.61	3100 years[†]
Polar ice, glaciers	29,000	2.08	16,000 years
Ground water (actively exchanged)[‡]	4,000	0.295	300 years
Freshwater lakes	125	0.009	1–100 years[§]
Saline lakes	104	0.008	10–1000 years[§]
Soil and subsoil moisture	67	0.005	280 days
Rivers	1.2	0.00009	12–20 days[¶]
Atmospheric water vapor	14	0.0009	9 days

[*] Modified from R. G. Wetzel (1983, 1999) after Kalinin and Bykov. Slightly different values are given by Shiklomanov, but ratios are similar.
[†] Based on net evaporation from the oceans.
[‡] Kalinin and Bykov estimated that the total ground water to a depth of 5 km in the Earth's crust amounts to 60 × 10^6 km^3. This is much greater than the U.S. Geological Survey estimate of 8.3 × 10^6 km^3 to a depth of 4 km. Only the volume of the upper ,actively exchanged ground water is included here.
[§] Renewal times for lakes vary directly with volume and mean depth, and inversely with rate of discharge. The absolute range for saline lakes is from days to thousands of years.
[¶] Twelve days for rivers with relatively small catchment areas of less than 100,000 km^2; 20 days for major rivers that drain directly to the sea.

productive wetlands, that is, areas of shallow water often found at the edge of lakes (littoral zone) or along rivers (floodplain). *See* ECOLOGY; LIMNOLOGY.

Ecosystem structure. Most water on the Earth's surface is saline in oceans (97.61%), and most of the remainder is polar ice (2.08%) and ground water (0.29%). Only 0.009% exists in freshwater ecosystems as temporary storage reservoirs on land with renewal times much shorter than those of the oceans (**Table 1**). The volume of water flowing from these reservoirs to the sea in rivers is very small (0.00009%) with a mean residence time of about 2 weeks. During high flow periods, surface water often recharges adjacent ground-water aquifers. Return flows from ground water to rivers occur during periods of low flow or drought, and usually maintain a base flow in river channels. *See* RIVER.

Water storage and retention times are altered by shifts in the balance between water inputs and losses. Lakes receive water from precipitation directly on the surface, from surface inflows from the drainage basin, and from subsurface ground-water seepage. Lakes lose water by flow from an outlet (drainage lakes), seepage through the basin walls into the ground water (seepage lakes), evaporation, and evapotranspiration from higher aquatic plants. Saline lakes occur in closed basins with no outflow except by evaporation.

Lakes are concentrated in the subarctic and temperate regions of the Northern Hemisphere, and reservoirs are concentrated in the subtemperate and subtropical zones. Although some 40% of the total volume of fresh water is contained in seven great lake basins of Siberia, North America, and eastern Africa, most lakes and reservoirs are very much smaller. Most of the millions of lakes and reservoirs are small and relatively shallow, usually less than 15 m (50 ft) in depth (**Fig. 1**). Mean depths are even shallower (less than 5 m or 17 ft), and as a result often over half of the sediments occur within the photic zone; therefore much of the sediments of lakes and reservoirs is potentially capable of supporting photosynthesis. Similarly, major flowing-water ecosystems occur in low-gradient tropical and subtropical regions, and they are posited to support high photosynthetic productivity in extensive floodplain and land-water interface regions. *See* LAKE; PHOTOSYNTHESIS.

The productivity and internal metabolism of aquatic ecosystems are driven and controlled by energy from solar radiation acquired by photosynthesis. Inland waters receive organic products of photosynthesis directly from their aquatic flora, and indirectly from their drainage basins as particulate and dissolved organic matter from terrestrial and wetland plants imported by stream water, storm runoff, and ground water.

In addition, solar energy is absorbed and dissipated as heat. As surface waters warm, the upper layers

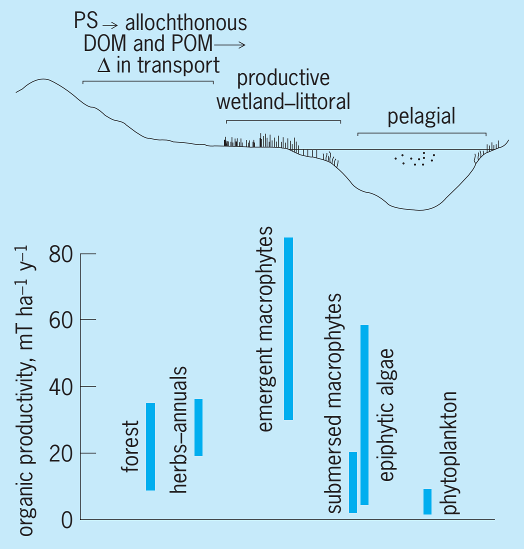


Fig. 1. Lake ecosystem showing the drainage basin with terrestrial photosynthesis (PS) of organic matter, movement of nutrients and dissolved (DOM) and particulate organic matter (POM) in surface and ground-water flows toward the lake basin, and chemical and biotic alteration of these materials en route, particularly as they pass through the highly productive and metabolically active wetland-littoral zone of the lake per se (net organic productivity in metric tons per hectare per year). (*Modified from R. G. Wetzel, 1983*)

TABLE 2. Mean composition of river waters of the world (mg L^{-1})*

	Ca^{+2}	Mg^{+}	Na^{+}	K^{+}	CO_3^{-2} (HCO_3^{-})	SO_4^{-2}	Cl^{-}	NO_3^{-}	Fe (as Fe_2O_3)	SiO_2	Sum
North America	21.0	5.	9.	1.4	68.	20.	8.	1.	0.16	9.	142
South America	7.2	1.5	4.	2.	31.	4.8	4.9	0.7	1.4	11.9	69
Europe	31.1	5.6	5.4	1.7	95.	24.	6.9	3.7	0.8	7.5	182
Asia	18.4	5.6	5.5	3.8	79.	8.4	8.7	0.7	0.01	11.7	142
Africa	12.5	3.8	11.	—	43.	13.5	12.1	0.8	1.3	23.2	121
Australia	3.9†	2.7	2.9	1.4	31.6	2.6	10.	0.05	0.3	3.9	59
World	15.	4.1	6.3	2.3	58.4	11.2	7.8	1.	0.67	13.1	120
Cations (μeq L^{-1})	750	342	274	59	—	—	—	—	—	—	1425
Anions	—	—	—	—	958	233	220	17	—	—	1428

*From R. G. Wetzel (1983, 1999) after data from several sources.
†Values of calcium are likely less, on the average, than sodium and magnesium in Australian surface waters.

are less dense and overlie deeper, cooler water (the maximum density of water occurs at a temperature of 4°C or 39°F; warmer water is less dense). This stratification generally persists until seasonal insolation (exposure to the Sun's rays) is reduced and water temperatures and densities of the different strata equalize. Density-induced stratification influences physical, chemical, and biological properties of lakes and reservoirs, including the thermal structure, water mass stratification, and hydrodynamic movements. Heat is distributed and altered by the physical work of wind energy, currents and other water movements, basin morphometry, and water losses. These characteristics of warmer, less dense water overlying cooler, more dense strata for long periods of time (months) structure the aquatic habitats and have marked attendant effects on all chemical cycles, metabolic rates, and population dynamics and their productivities.

The most abundant dissolved ions, constituting total salinity, in inland waters are usually four major cations [calcium (Ca^{+2}), magnesium (Mg^{+2}), sodium (Na^{+}), potassium (K^{+})] and four anions [bicarbonate (HCO_3^{-}), carbonate (CO_3^{-2}), sulfate (SO_4^{-2}), chloride (Cl^{-})]. The salinity of freshwaters has a world average concentration of about 120 mg/L, but varies among continents and with the lithology of land masses (**Table 2**).

The composition of cellular protoplasm of organisms is dominated by six elements (carbon, hydrogen, oxygen, nitrogen, phosphorus, sulfur). Biogeochemical cycles control the availability of phosphorus, nitrogen, and several minor nutrients, and these elements often limit biotic development. Phosphorus, in comparison with other macronutrients required by biota, is least abundant and commonly the first element to limit biological productivity. When phosphorus is in adequate supply, nitrogen availability invariably limits productivity (**Table 3**). Standing inland waters have been categorized into

TABLE 3. General ranges of primary productivity of phytoplankton and related characteristics of lakes of different trophic categories*

Trophic type	Mean primary productivity, mg C m^{-2} day^{-1}†	Phytoplankton density, cm^3 m^{-3}	Phytoplankton biomass, mg C m^{-3}	Chlorophyll a, mg m^{-3}	Dominant phytoplankton	Light extinction coefficents, η m^{-1}	Total organic carbon, mg L^{-1}	Total P, μg L^{-1}	Total N, μg L^{-1}	Total inorganic solids, mg L^{-1}
Ultraoligotrophic	<50	<1	<50	0.01–0.5		0.03–0.8		<1–5	<1–250	2–15
Oligotrophic	50–300		20–100	0.3–3	Chrysophyceae Cryptophyceae	0.05–1.0	<1–3			
Oliogomesotrophic		1–3			Dinophyceae Bacillariophyceae			5–10	250–600	10–200
Mesotrophic	250–1000		100–300	2–15		0.1–2.0	<1–5			
Mesoeutrophic		3–5						10–30	500–1100	100–500
Eutrophic	>1000		>300	10–500	Bacillariophyceae Cyanobacteria	0.5–4.0	5–30			
Hypereutrophic		>10			Chlorophyceae Euglenophyceae			30 to >5000	500 to >15,000	400–60,000
Dystrophic	<50–500		<50–200	0.1–10		1.0–4.0	3–30	<1–10	<1–500	5–200

*From R. G. Wetzel (1983, 1999) after many authors and sources.
†Referring to approximately net primary productivity, such as measured by the ^{14}C method.

various trophic scales on the basis of ranges of major nutrients and algal productivity (Table 3). *See* BIOGEOCHEMISTRY; BIOLOGICAL PRODUCTIVITY.

Organic matter is either soluble or particulate; once organic matter enters an aquatic ecosystem, it tends to remain there. Oxygen, however, is quite insoluble in water, and most oxygen produced in an aquatic system will tend to diffuse to the atmosphere. When organic matter in lakes is utilized and biochemically respired by heterotrophic organisms, the amount of organic matter to be oxidized can be much larger than the oxygen available for oxidation. In large lakes, inputs of organic matter from internal photosynthesis and from external sources are usually small relative to the large amount of dissolved oxygen in the water. In contrast, small lakes have a greater proportion of water supporting photosynthesis in wetland and littoral (pertaining to shores of lakes, seas, or oceans) areas, import proportionally more organic materials from terrestrial habitats, and have a smaller pool of dissolved oxygen. Thus, respiration can exhaust oxygen dissolved in the lake water in lower strata that are poorly connected hydrodynamically with the atmosphere. The exhaustion of dissolved oxygen in large portions of lake ecosystems is exacerbated by the process of eutrophication (the increased photosynthetic production of organic matter in response to excessive loading of nutrients, particularly phosphorus).

Once the dissolved oxygen is consumed, other dissolved ions [nitrate (NO_3^{-1}), sulfate (SO_4^{-2}), iron (Fe^{+2}), manganese (Mn^{+2}), and organic compounds] can function as alternative electron acceptors under anoxic conditions. These processes are less efficient energetically than oxidative decomposition, and the rates and efficacy of decomposition decline under anaerobic conditions, such as in anoxic lower strata of productive lakes.

Diversity of habitats. Freshwater ecosystems consist of entire drainage basins. The nutrient and organic matter content of drainage water from the catchment area is modified as water moves downgradient to and within the lake or reservoir per se (**Fig. 2**). Autotrophic photosynthetic productivity is generally low to intermediate in the terrestrial components, highest in the wetland interface region between the land and water, and lowest in the open water (Fig. 2). Along the gradient from land to river channels, the greatest productivity is also in the marginal floodplain regions. Autotrophic productivity in river channels is generally low as in the pelagic (open water, nonlittoral) regions of lakes. Most of the organic matter utilized in running waters is imported from floodplain and terrestrial sources.

The land-water interface region of aquatic ecosystems is always the most productive per unit area along the gradient from land to open water of both lakes and reservoirs. Because most aquatic ecosystems occur in geomorphologically mature terrain of gentle slopes and are small and shallow, the wetland-littoral components usually dominate in productivity and in synthesis of organic matter. The region of greatest productivity is the emergent macrophyte zone. Emergent aquatic plants have structural and physiological adaptations that not only tolerate the hostile reducing anaerobic sediments but exploit the high nutrient and water availability of this habitat. Nutrients entering the zone of emergent macrophytes tend to be assimilated by microflora of the sediments and detritus particles, and are then recycled to the emergent macrophytes. Export from the emergent zone is dominated by dissolved organic compounds released from decomposition of plant detrital material. Nutrients, particularly nitrogen and phosphorus, tend to be conserved and retained in the wetland-littoral regions.

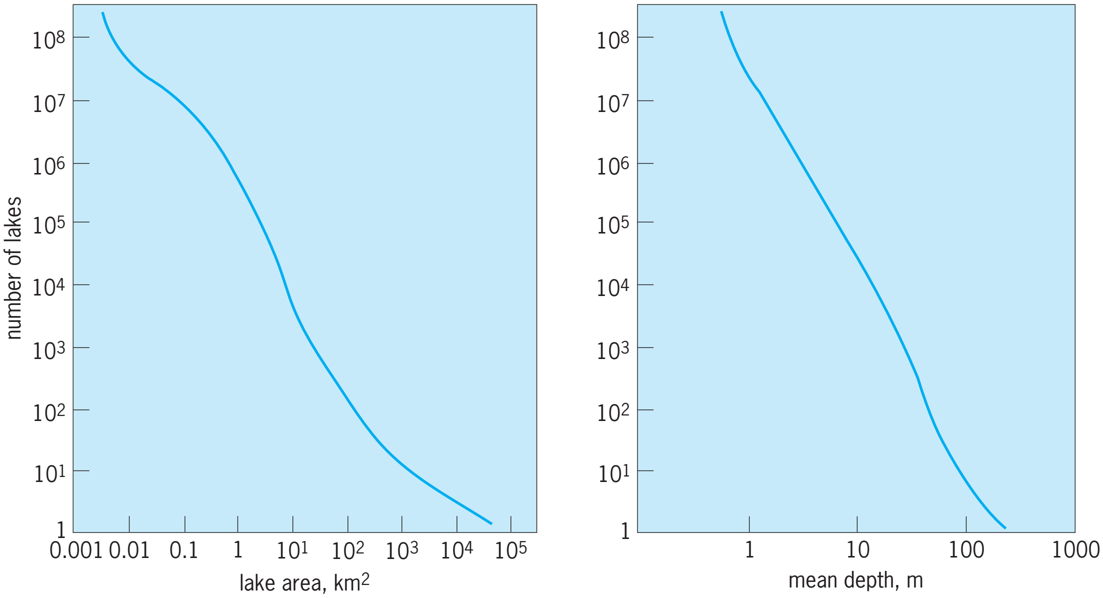


Fig. 2. Approximate number of lakes of the world in relation to lake area and approximate mean water depth. (*After R. G. Wetzel, 1990*)

Of the total light energy entering water, a portion is scattered, and the remainder is exponentially attenuated by absorption of the water molecules themselves, dissolved organic compounds (particularly humic substances), and suspended particulate matter. Submersed macrophytes are limited physiologically by reduced availability of light underwater and by slow diffusion rates of gases and nutrients in water. Internal recycling of resources, both of gases [carbon dioxide (CO_2), oxygen (O_2)] of metabolism and of nutrients, are important to the abilities of submersed plants to function and grow as well as they do under light and gas limitations.

The second most productive component of the wetland-littoral community is the microflora attached to aquatic plants and other surfaces. The surfaces provided by aquatic plants in lakes and rivers can be extensive, often exceeding 25 m^2 per square meter of bottom. The physiology and growth of attached microflora are intimately coupled to the physical and metabolic dynamics of the living substrata upon which they grow. High sustained growth of attached microflora results from their recycling of essential gases (CO_2, O_2) and dissolved nutrients. Nutrient uptake is directed primarily to the high net growth of attached microflora, and it is responsible for the high capacity of wetland-littoral areas to improve the quality of water passing through these communities.

The wetland-littoral complex, including the marginal floodplains of many rivers, produces the major sources of organic matter and energy of many fresh-water ecosystems. Most of the particulate organic matter is decomposed within these interface regions. Organic matter is exported predominantly as dissolved organic matter to the recipient lake or river (**Fig. 3**).

The deep-water pelagic zone of lakes is the least productive region along the gradient from land to water, regardless of nutrient availability. Growth of phytoplanktonic algae of the pelagic zone is limited by sparse distribution in a dilute environment where nutrient recycling is restricted by the sinking of senescent phytoplankton below the depth of photosynthesis. When nutrient recycling and availability is increased, greater phytoplankton cell densities attenuate underwater light and reduce the volume of water in which photosynthesis occurs. Despite low productivity per unit area, pelagic productivity can be collectively important in large lakes and for higher trophic levels that depend upon this organic matter.

Among biological communities, functionally similar organisms can be grouped into trophic levels based on similarities in patterns of organic matter production by photosynthesis and consumption by heterotrophic microbes and animals. Energy is transferred and nutrients are cycled within an overall ecosystem trophic structure. The productivity of each trophic level is the rate at which energy enters the trophic level from the next lower level. Because organisms expend considerable energy for maintenance and since death of an organism routes much energy and nutrients into the detrital pool of dead organic matter, only a portion of the energy of one trophic level (5–15%) is available for transfer and use by higher trophic levels. Available energy and efficiency of energy transfer decreases progressively at higher trophic levels, so that rarely can more than five or six trophic levels be supported.

A second trophic level consists of zooplankton (dominated by four major groups of animals: Protozoa/Protista, Rotifera, and the crustaceans Cladocera and Copepoda) and benthic invertebrates. In the pelagic zone these herbivorous organisms are consumed by small fishes, fry of larger fishes, and predatory zooplankton, which constitute a third trophic level (primary carnivores). A fourth trophic level may consist of medium-sized piscivorous fishes; and a fifth, large piscivorous fishes. Higher trophic levels are rare in freshwater ecosystems.

The species composition of the higher trophic levels affects the pathways of energy utilization from lower trophic levels. For example, efficiency

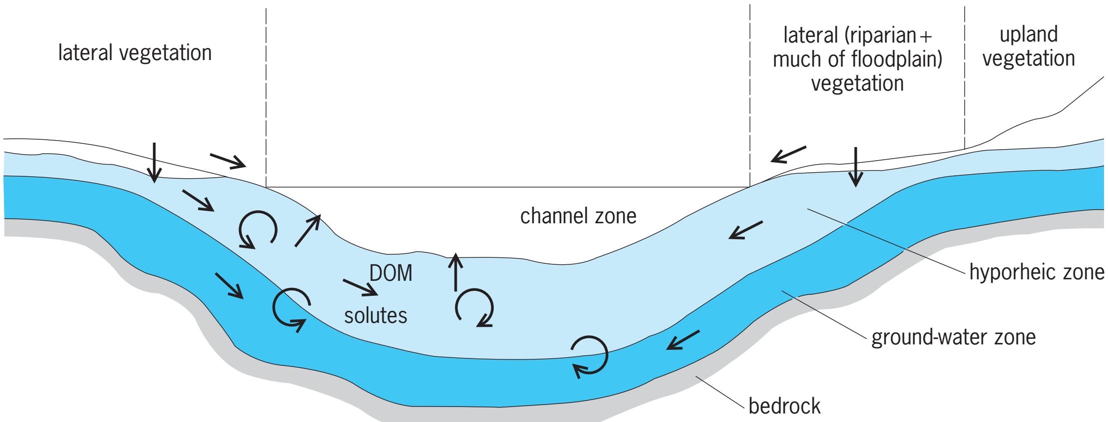


Fig. 3. Lateral and vertical boundaries of flowing-water ecosystems. The stream ecosystem boundary is defined as the hyporheic–ground-water interface and thereby includes a substantial volume beneath and lateral to the main channel. Vegetation rooted in the hyporheic zone is therefore part of stream ecosystem production. Arrows indicate flow pathways of dissolved organic matter and inorganic solutes derived from plant detritus within the stream ecosystem. (*Modified from R. G. Wetzel and A. K. Ward, 1992*)

of consumption of primary production by zooplankton is often appreciably greater in the absence of zooplankton-feeding fishes than in their presence. The community structure of phytoplankton responds variably to grazing impacts in concert with their available resources (light, nutrients, organic constituents) and may or may not be able to compensate for grazing losses in overall primary production. *See* COPEPODA; POPULATION ECOLOGY; ROTIFERA; ZOOPLANKTON.

Interactive regulatory mechanisms. Most of the organic carbon of aquatic ecosystems, and hence energy available for system operation, exists as dead organic matter (detritus), in both dissolved and particulate forms. Dissolved and particulate organic carbon moves with the water; residual particulate organic carbon is ultimately deposited at the bottom of static water. Dissolved organic carbon can also sediment if adsorbed to particulate matter or if polymerization occurs. Most heterotrophic decomposition occurs in benthic regions where organic carbon is often largely displaced from sites of production to the dominant benthic sites of decomposition at the sediment-water interface.

Although the specific composition of organic matter varies greatly, detritus inevitably carries most of the energy of the ecosystem from its points of photosynthetic origins to places of transformation. Most of that heterotrophic transformation does not occur by digestion in predatory metazoan animals, but by microbes and often under anaerobic conditions in sediments. Productivity of microbes is maximized within the resources available from dead organic matter and nutrients. Microbial heterotrophic utilization is highly dynamic and changes with sufficient rapidity to allow many generation turnovers before predators with much slower turnover rates can respond reproductively. Viral mortality of microbes as well as ingestion of bacteria by protists (heterotrophic flagellates and protozoa of the microbial loop) diverts much organic carbon from animal trophic levels. That productivity, much of which is not ingested by higher metazoan animals, is of major importance and usually dominates the energy flows within the pelagic region and completely dominates energy flows of the complete fresh-water ecosystems (wetlands, littoral, benthic, and pelagic compartments). Microbial heterotrophy is of importance to higher trophic levels in feedback processes, both positively (such as nutrient recycling and utilization by primary producers) and negatively (such as oxygen consumption and production of toxic fermentative metabolic end products). The abundance, distribution, and microbial decomposition of dissolved and particulate detritus both regulate and stabilize energy metabolism and nutrient availability in aquatic ecosystems. *See* FOOD WEB.

Relation to society. Freshwater is the pivotal resource upon which agriculture, industry, and domestic life depend. Natural or artificially created basins contain water for various periods of time. These bodies of water function in multiple ways in human activities—as sources of water, sources of biota for nutrition, and sources of recreational and esthetic value to human well-being. These water bodies are also used as purification systems, intentionally or fortuitously, for various waste products of human activities. Many of the purposes and uses to which lakes, reservoirs, and rivers are put are incompatible.

Nearly all problems of alteration of surface waters and ground-water aquifers focus on (1) introduction of abnormal amounts of alien substances or biota, or (2) modification of retention or distribution of the water by excessive use. Eutrophication of surface water results from loading of nutrients, particularly phosphorus, to lakes and reservoirs which leads to excessive production of certain algae and higher aquatic plants, and often a decrease in biodiversity. Excessive atmospheric loadings of sulfur and nitrogen oxides from fossil fuel combustion have resulted in high loadings of strong acids in liquid (rainfall) and particulate (dust) precipitation. The resulting acidification of surface and ground waters has radically altered the water chemistry and biota, with a marked reduction in biodiversity. Pollutants, such as heavy metals (mercury, copper, lead) and toxic organic compounds, continue to contaminate many surface and ground waters to the extent that not only has biodiversity been reduced but many of these water reservoirs cannot be used for human activities. In some cases, introduction of exotic species has similarly resulted in competitive exclusion of many endemic species and has also altered chemical conditions of the water. In all of these cases of environmental alterations, paleolimnological analyses of the mineralogy and structure of sediments, their organic and inorganic chemical constituents, and the fossil remains of organisms preserved in the sediments permit interpretations about past states and changing conditions of the ecosystem.

A great diversity of terrain, geomorphology, climate, and hydrology occurs in every country. This natural diversity is overlain with cultural histories in which humans have systematically and often ingeniously modified hydrological patterns, particularly of surface waters, to accommodate agricultural and water supply needs. Numerous political and human conflicts originated or were incited by the need for or mismanagement of freshwaters. The present global warming associated with increased loading of carbon dioxide and other pollutant gases to the atmosphere will certainly alter regional climatic patterns, hydrological regimes, and the availability and quality of fresh-water ecosystems. *See* ACID RAIN; EUTROPHICATION; WATER POLLUTION. Robert G. Wetzel

Bibliography. P. L. Brezonik, *Chemical Kinetics and Process Dynamics in Aquatic Systems*, Lewis Publishers, Boca Raton, FL, 1994; G. A. Cole, *Textbook of Limnology*, 3d ed., 1988; G. D. Cooke et al., *Restoration and Management of Lakes and Reservoirs*, Lewis Publishers, Boca Raton, FL, 1993; P. H. Gleick (ed.), *Water in Crisis: A Guide to the World's Fresh Water Resources*, Oxford University Press, New York, 1993; D. Gunnison (ed.), *Microbial Processes in Reservoirs*, 1985; G. E. Hutchinson, A *Treatise on Limnology: I. Geography, Physics, and Chemistry*

(1957), *II. Introduction to Lake Biology and the Limnoplankton* (1967), *III. Limnological Botany* (1975), *IV. The Zoobenthos* (1994), John Wiley, New York; O. T. Lind, *Handbook of Common Methods in Limnology*, 2d ed., 1985; H. Loffler, *Paleolimnology*, 1987; M. Straskraba and A. H. Gnouck, *Freshwater Ecosystems*: *Modelling and simulation*, Develop. Environ. Model. 8, 1985; R. Thomas et al. (eds.), *Ecological Effects of In-Situ Sediment Contaminants*, 1987; R. G. Wetzel, Death, detritus, and energy flow in aquatic ecosystems, *Freshwater Biol.*, 33:83–89, 1995; R. G. Wetzel, *Limnology: Lake and River Ecosystems*, 3d ed., Academic Press, San Diego, 2000; W. D. Williams, *Life in Inland Waters*, 1983.

Friction

Resistance to sliding, a property of the interface between two solid bodies in contact. Many everyday activities like walking or gripping objects are carried out through friction, and most people have experienced the problems that arise when there is too little friction and conditions are slippery. However, friction is a serious nuisance in devices that move continuously, like electric motors or railroad trains, since it constitutes a dissipation of energy, and a considerable proportion of all the energy generated by humans is wasted in this way. Most of this energy loss appears as heat, while a small proportion induces loss of material from the sliding surfaces, and this eventually leads to further waste, namely, to the wearing out of the whole mechanism. *See* WEAR; WORK.

In stationary systems, friction manifests itself as a force equal and opposite to the shear force applied to the interface. Thus, as in **Fig. 1**, if a small force S is applied, a friction force P will be generated, equal and opposite to S, so that the surfaces remain at rest. P can take on any magnitude up to a limiting value F, and can therefore prevent sliding whenever S is less than F. If the shear force S exceeds F, slipping occurs. During sliding, the friction force remains approximately equal to F (though often it is smaller by about 20%) and always acts in a direction opposing the relative motion. The friction force is proportional to the normal force L, and the constant of proportionality is defined as the friction coefficient f. This is expressed by the equation $F = fL$.

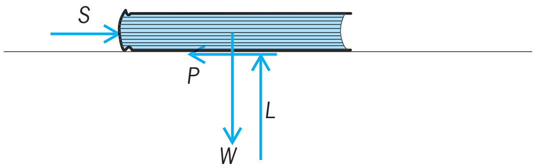


Fig. 1. The forces acting on a book resting on a flat surface when a sheer force *S* is applied. The friction from *P* is equal to *S* (up to a limiting value *F*), while *L*, the normal force, is equal to the weight *W* of the book.

The "laws" of friction are essentially statements about the friction coefficient, and have been worked out over the past five centuries by a number of distinguished engineers and scientists, among them Leonardo da Vinci and Charles de Coulomb. These laws, which are approximately, but never perfectly, obeyed by typical sliding systems are as follows: (1) The friction coefficient is independent of the normal force. This is another way of saying that the friction force is proportional to the normal force. (2) The friction coefficient is independent of the sliding velocity. However, as was stated above, the friction for surfaces at rest is often about 20% greater than for the same system when sliding. Such systems are often spoken of as having larger static friction than kinetic friction. (3) The friction coefficient is independent of the apparent area of contact. Thus, flat surfaces and surfaces with contacting asperities but of the same projected area give the same friction coefficient. (4) The friction coefficient is independent of surface roughness.

Actually, the friction coefficient is primarily a property of the contacting materials and the contaminants or lubricants at the interface.

Closely related to the friction coefficient is the concept of friction angle. The friction angle θf is the largest angle relative to the horizontal at which a surface may be tilted, so that an object that is placed on the surface does not slide down. Leonhard Euler was the first to show that tan θ_f is equal to the friction coefficient.

Mechanism. A number of processes occur at the interface between two solids that tend to inhibit sliding or to use up energy during sliding. All these contribute to friction. Over the last three centuries, there has been much discussion over which one of these processes is the most important and can thus be considered the main cause of friction. Until about 1940 it was considered that surface roughness was the main cause of friction, and that work had to be done during sliding to lift one sliding surface over the high spots on the other one. Modern work has largely discounted the importance of surface roughness: first, from the standpoint of theoretical considerations, because the work done in sliding up a high spot on a surface is largely recovered on the down side, and second, because experimental testing shows that for most sliding systems friction coefficients are largely independent of the roughness of the surfaces. Very smooth surfaces, like cleaved mica, which is smooth to within one atomic diameter, give friction at least as great as that of ordinary surfaces. For reasons that are not clear, most people find it very hard to accept the fact that smooth surfaces give as much friction as rough ones.

The modern theory of friction attributes friction to adhesion between surface atoms. These surface atoms form bonds, similar in strength to those that hold the solids themselves together, and friction represents the force required to break the bonds. Atomic forces are known to be of very short range (about 10^{-10} m), and accordingly the bonds are appreciable only at those small patches (called junctions) at which the surface atoms come within 10^{-10} m of each other. All the friction originates at these junctions, which make up the real area of contact. Over the rest of the apparent area of contact the separation of the surfaces is generally many hundreds or thousands of atomic diameters, and essentially none of the friction originates there. In a typical sliding system involving relatively large surfaces, the real area of contact is less than 1% of the apparent area.

According to the adhesion theory, lubricants at the surface reduce friction because they form bonds of reduced strength. Since hard solids generally form strong bonds while liquids and soft solids form weak bonds, it comes as no surprise to discover that good lubricants are generally either liquids or soft solids.

Besides the adhesion mechanism, which generally accounts for more than 90% of the total friction force, there are a few minor mechanisms that act so as to increase the friction. One is a plowing mechanism, when a hard surface slides against a soft one, and plows out a series of grooves. A second one, with very rough surfaces, consists of a component of the motion perpendicular to the interface. (This is the roughness effect, which at one time was thought to be the main component of friction.) Third, there is elastic hysteresis, as result of the fact that there is elastic compression and then relaxation of the material near the contacting interface, and not all the energy is recoverable. Fourth, there is an electrostatic attraction between the surfaces, but this is only significant when electrically insulating materials are used.

Almost all the frictional energy appears as heat at the interface between the sliding surfaces. This frictional heat was used by humans in prehistoric times to light fires, and this use survives today in the striking of matches, with chemical combustion initiated by the temperature rise produced by sliding. However, frictional heat usually is a nuisance, and sliding surfaces must often be cooled to prevent heat damage.

Friction coefficient values. Values of the friction coefficient are required in analyzing many problems in mechanics, in order to estimate the frictional forces at interfaces and to compute their consequences. **Figure 2** shows typical friction coefficient values, plotted as a function of the state of cleanliness or of lubrication of the surfaces. As will be seen, friction coefficients range from about 1.5 to about .07 depending on the sliding conditions. Unlubricated metals give higher friction than nonmetals, but well-lubricated metals give less. In consequence, when sliding mechanisms giving low friction are desired, metals are used when a lubricant is available, but nonmetals are preferred when the surfaces must operate in an unlubricated condition.

It should be emphasized that the friction coefficient values shown in Fig. 2 are typical ones, and that in any specific case friction values differing by as much as 30% are quite likely to be found. Indeed, it is not yet possible, even for an expert in the friction field, to be able to estimate the friction coefficient in any given case to within 10% or better.

A number of sliding systems are exceptional, in that they give friction values that are rather different from those shown in Fig. 2, and much use is made of their exceptional friction properties. Elastomeric materials like rubber give friction values that are about twice as great as those of other nonmetals, and this explains the use of rubber in shoe soles and heels, in automobile tires, and in other cases in which good traction is required. Solids that give exceptionally low friction, about half the values for other nonmetals, include graphite, Teflon, and ice. Teflon and graphite are used in low-friction coatings, while the low friction of ice is made use of in skating and skiing. Perhaps the lowest friction values are found in mammalian joints (like the human knee joint), where it has been found that friction coefficient values are generally as low as .02. It is not clear how such a low value is achieved.

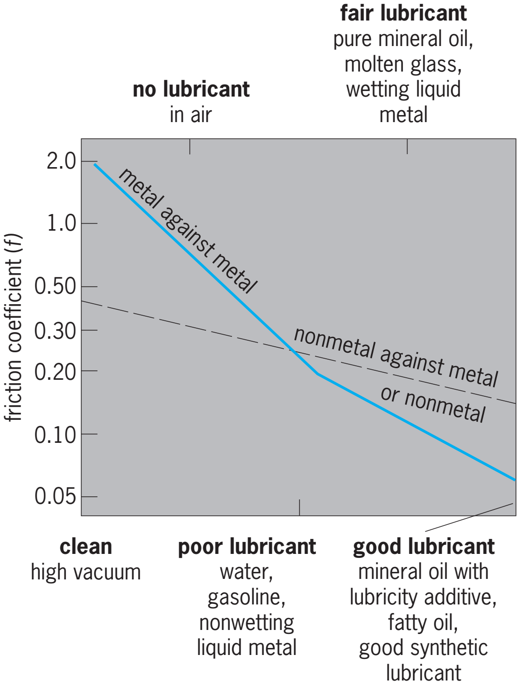


Fig. 2. Typical friction coefficient values.

Two special cases of low friction should be mentioned. One involves the use of rolling rather than sliding action. This generally produces much lower friction than is possible during sliding. For example, ball bearings generally give friction coefficients in the range .002 to .005. The other involves fluid-lubricated sliding systems at high sliding speeds, in which hydrodynamic effect allows separation of the surfaces by a full fluid film. Here friction coefficient values of .001–.003 are common.

Stopping distances. For a sliding object moving with velocity v, the distance required to bring it to a stop through friction action is governed by the equation $s = v^2/2gf$, where g is the acceleration of gravity. The distance is independent of the weight of the object. The typical stopping distances given in the **table** assume an initial speed of 60 mi/h (100 km/h). The distances are minimal, assuming in the case of rolling vehicles that the brakes are applied severely enough to cause locking of the wheels, or a braking action just short of that.

Stopping distances at 60 mi/h (100 km/h)

Typical system	Friction coefficient	Stopping distance, ft (m)
Car on dry road	.8	151 (46)
Car on wet road	.5	242 (74)
Train on dry rails	.3	403 (123)
Train on wet rails	.2	605 (184)
Car on icy road	.10	1210 (369)

Many traffic accidents at level railroad crossings occur because drivers of cars are not aware of the great length of stopping distances for trains. Road accidents frequently take place because drivers used to high friction coefficient values under dry road conditions do not realize that weather conditions have changed and that lower friction values of wet surfaces apply.

Human interactions with friction. For normal walking, humans require friction coefficient values of .20 or above. When surfaces with lower friction coefficient values are encountered, people instinctively slow down, shorten their stride, and are able to walk satisfactorily when the friction is as low as .10. Slipping generally occurs when someone walking normally on a high friction surface encounters a low-friction spot (like a banana peel), and there is no opportunity to adjust stride length.

When lifting an object by hand, a neurophysiological mechanism comes into play so that the object is gripped with a force that is only a little larger than would allow it to slip.

Frictional oscillations. It was noted above that the laws of friction are not perfectly obeyed. This generally means that if one of the variables (for example, load or apparent area) is varied by a factor of 10, the friction coefficient is changed by less than 10%. Only one of these departures from constancy has practical significance, namely, if the friction goes down when the sliding speed goes up (a so-called negative characteristic), frictional vibrations may be produced. At slow sliding speeds the vibrations take the form of relaxation oscillations (creaking, chattering, and groaning noises), while at high speeds they are harmonic oscillations (squeaking). As a general rule, unlubricated and poorly lubricated surfaces give a negative characteristic and show a tendency toward oscillation, while well-lubricated surfaces give a positive characteristic and are immune to oscillations.

Friction oscillations are generally referred to as “stick-slip,” but strictly speaking this term applies only to relaxation oscillations, because during part of each oscillation cycle the sliding surfaces actually stick and have zero relative velocity. There is always some interfacial slip during harmonic oscillations.

Frictional oscillations can generally be avoided by changing one or more of the parameters. Helpful changes include reducing the normal load, increasing the sliding speed, introducing damping into the system, stiffening all components acting as springs, changing materials, or introducing good lubricants.

Frictional oscillations can be useful. They are the mechanism of music production in bowed stringed instruments, and they have been used occasionally to detect the entry of intruders into a home. More significantly, in many sliding systems they act as a warning signal that the lubricant has disappeared, that excessive wear damage is imminent, and that relubrication is in order.

Significant friction problems. In prehistoric and early historic times, humans’ main interest in friction was to reduce the friction coefficient, to reduce the labor involved in dragging heavy objects. This led to the invention of lubricants, the first of which were animal fats and vegetable oils. A great breakthrough was the use of rolling action—in the form of rolling logs and then wheels—to take advantage of the lower friction coefficients of rolling systems.

In modern engineering practice available materials and lubricants reduce friction to acceptable values. In special circumstances when energy is critical, determined efforts to minimize friction are undertaken. Friction problems of practical importance are those of getting constant friction in brakes and clutches, so that jerky motion is avoided, and avoiding low friction in special circumstances, such as when driving a car on ice or on a very wet road. Also, there is considerable interest in developing new bearing materials and new lubricants that will produce low friction even at high interfacial temperatures and maintain these properties for long periods of time, thus reducing maintenance expenses. Perhaps the most persistent problem is that of avoiding frictional oscillations, a constant cause of noise pollution of the environment. *See* LUBRICANT.

Ernest Rabinowicz

Bibliography. American Society for Metals, *Friction, Lubrication, and Wear Technology*, 1992; F. P. Bowden and D. Tabor, *Friction: An Introduction to Tribology*, 1973 reprint 1982; M. Grunze and H. J. Kreuzer (eds.), *Adhesion and Friction*, 1989; I. M. Hutchings, *Tribology: Friction and Wear of Engineering Materials*, 1989.

Friedel-Crafts reaction

Originally a substitution reaction, catalyzed by aluminum chloride, in which an alkyl (R—) group or an acyl (RCO—) group replaces a hydrogen atom of an aromatic nucleus. Later, the term came to imply any of a group of electrophilic reactions in which the production of an intermediate cation or positively polarized complex species was catalyzed, either directly or indirectly, by the presence of a Lewis acid or a Brønsted-Lowry acid. The broadened definition embraces some addition, isomerization, elimination, and cracking reactions as well as the synthetically important alkylation or acylation of aromatic systems. *See* ACID AND BASE.

Alkylation. In the classical alkylation reaction, an alkyl halide (RX) serves as the alkylating agent. In reaction (1), an excess of the hydrocarbon benzene

$$\mathrm{R{-}Cl} + \mathrm{C_6H_6} \xrightarrow{\mathrm{AlCl_3}} \mathrm{R{-}C_6H_6} + \mathrm{HCl} \qquad (1)$$

(C_6H_6) is the solvent. In case the aromatic hydrocarbon is not so readily available, a smaller quantity may be used in an inert solvent such as carbon disulfide or petroleum ether.

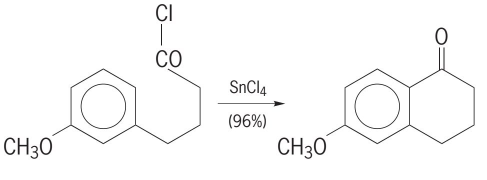
Alkylations usually need less than 1 mole of aluminum chloride ($AlCl_3$) per mole of alkyl halide, and a molar ratio of 1:10 is common. As might be predicted from the relative ease of ionization, tertiary halides (R_3CX) are superior to secondary (R_2CHX) or primary (RCH_2X) halides for the alkylation reaction. As a consequence of the possible rearrangement of the carbocationic intermediate, the alkyl group which becomes attached to the aromatic ring may differ in structure from that of the alkyl halide employed. For example, with *n*-propyl chloride ($CH_3CH_2CH_2Cl$), a large part of the product would be isopropylbenzene, $C_6H_5CH(CH_3)_2$, rather than the expected *n*-propylbenzene, $C_6H_5CH_2CH_2CH_3$. Since alkyl groups attached to an aromatic nucleus facilitate further substitution, it is difficult to prevent di- or polysubstitution during the reaction. The structure of the dialkylation and trialkylation products cannot be predicted simply from a knowledge of the rules of substitution, since aluminum chloride is a catalyst for the intermolecular and intramolecular migration of alkyl groups. Alkenes (C_nH_{2n}) may be substituted for alkyl halides in the Friedel-Crafts reaction, and the industrial synthesis of ethylbenzene and isopropylbenzene (cumene) has been carried out in this way.

Acylation. For acylation of aromatic hydrocarbons, acyl halides [reaction (2)] have proved most valuable although acid anhydrides [reaction (3)] have also been used. For each mole of acid chloride present, slightly more than 1 mole of $AlCl_3$ is commonly required, whereas with acid anhydrides slightly more than 2 moles of $AlCl_3$ are necessary. The acyl group deactivates the benzene ring toward further substitution. Therefore, it is usually easy to obtain good yields of monosubstitution products. In contrast to the alkyl group, the substituted acyl group does not undergo rearrangement, making acylation a more reliable synthetic tool than the alkylation reaction.

$$R{-}COCl + C_6H_6 \xrightarrow{AlCl_3} R{-}COC_6H_6 + HCl \qquad (2)$$

$$(RCO)_2O + C_6H_6 \xrightarrow{AlCl_3} RCOC_6H_5 + RCOOAlCl_2 + HCl \qquad (3)$$

An important application of the Friedel-Crafts acylation has been in cyclizations via an intramolecular reaction. Reaction (4), the cyclization of 4–(3-methoxyphenyl)butanoyl chloride to yield 6-methoxytetralone, is an example. *See* ACYLATION.



(4)

Catalysts. In addition to aluminum halides, a wide variety of Lewis acids, chiefly acidic halides, have been used as catalysts. Some important examples are boron trifluoride, zinc chloride, stannic chloride, and antimony pentafluoride; arenetricarbonylmolybdenum has also been recommended as a Friedel-Crafts catalyst. Of the Brønsted type, H_2SO_4, HF, and trifluoroacetic acid are important, but much the same function can be served by solid acids, such as acidic cation exchange resins or aluminum silicates. *See* AROMATIC HYDROCARBON; BENZENE; CATALYSIS; SUBSTITUTION REACTION.

Charles K. Bradsher

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 10, 1994; G. A. Olah, *Friedel-Crafts Chemistry*, 1973.

Fringe (optics)

The part of optics using the light or dark bands produced by interference or diffraction of light. Distances between fringes are usually very small, because of the short wavelength of light. Fringes are clearer and more numerous when produced with light of a single color.

Diffraction fringes are formed when light from a point source, or from a narrow slit, passes by an opaque object of any shape. The Fraunhofer diffraction fringes produced when the light approaches and leaves the obstacle in essentially plane waves are especially important in the theory of optical instruments. *See* DIFFRACTION; RESOLVING POWER (OPTICS).

Interference fringes are obtained by bringing together two or more beams of light that have originated from a common source. This is usually accomplished by means of an apparatus especially designed for the purpose called an interferometer, although interference fringes may also be seen in nature. Examples are the colors in a soap film and in an oil film on water. When the fringes are controllable, for example, by changing the paths traversed by the two beams in an interferometer, they are valuable for accurate measurement of small distances and of slight differences in refractive index. *See* INTERFERENCE OF WAVES; INTERFEROMETRY.

If the light from a laser is passed through an interferometer, features such as coherence and phase fluctuations may be checked by observing the sharpness and drift of the fringes. *See* LASER.

F. A. Jenkins; W. W. Watson

Front

An elongated, sloping zone in the troposphere, within which changes of temperature and wind velocity are large compared to changes outside the zone. Thus the passage of a front at a fixed location is marked by rather sudden changes in temperature and wind and also by rapid variations in other weather elements such as moisture and sky condition. *See* TROPOSPHERE.

In its idealized sense, a front can be regarded as a

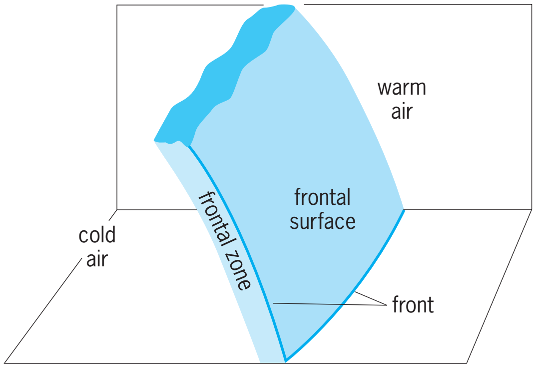


Fig. 1. Schematic diagram of the frontal zone, the angle with Earth's surface much exaggerated.

sloping surface of discontinuity separating air masses of different density or temperature. In practice, the temperature change from warm to cold air occurs mainly within a zone of finite width, called a transition or frontal zone. The three-dimensional structure of the frontal zone is illustrated in **Fig. 1**. In typical cases, the zone is about 0.6 mi (1 km) in depth and 60–120 mi (100–200 km) in width, with a slope of approximately 1/100. The cold air lies beneath the warm in the form of a shallow wedge. Temperature contrasts generally are strongest at or near the Earth's surface, the frontal zone usually being narrowest near the ground and becoming wider and more diffuse with height. Frontal zones seldom extend more than several miles above the Earth's surface. There are, however, sharp, narrow fronts of limited extent often present in the middle and upper troposphere, as discussed at the end of this article. With the exception of that section, the following concerns fronts at the Earth's surface.

The surface separating the frontal zone from the adjacent warm air mass is referred to as the frontal surface, and it is the line of intersection of this surface with a second surface, usually horizontal or vertical, that strictly speaking constitutes the front. According to this more precise definition, the front represents a discontinuity in temperature gradient rather than in temperature itself. The boundary on the cold air side is often ill-defined, especially near the Earth's surface, and for this reason is not represented in routine analysis of weather maps. In typical cases about one-third of the temperature difference between the Equator and the pole is contained within the narrow frontal zone, the remainder being distributed within the warm and cold air masses on either side. *See* AIR MASS; WEATHER MAP.

The wind gradient, or shear, like the temperature gradient, is larger within the frontal zone than on either side of it. In well-developed fronts the shift in wind direction often is concentrated along the frontal surface, while a more gradual change in wind speed may occur throughout the frontal zone. An upper-level jet stream normally is situated above a well-defined frontal zone. *See* JET STREAM.

A front moves approximately with the speed of the wind component normal to it. The strength of this component varies with season, location, and individual situation but generally lies in the range of 0–50 mi/h (0–80 km/h); 25 mi/h (40 km/h) is a typical frontal speed.

Frontal waves. Many extratropical cyclones (the lows on a weather map) begin as wavelike perturbations of a preexisting frontal surface. Such cyclones are referred to as wave cyclones. The life cycle of the wave cyclone is illustrated in **Fig. 2**. In stage I, prior to the development, the front is gently curved and more or less stationary. In stage II the front undergoes a wavelike deformation, the cold air advancing to the left of the wave crest and the warm air to the right. Simultaneously a center of low pressure and of counterclockwise wind circulation appears at the crest. The portion of the front which marks the leading edge of the cold air is called the cold front. The term warm front is applied to the forward boundary of the warm air. During stage III the wave grows in amplitude and the warm sector narrows. In the final stage the cold front overtakes and merges with the warm front, forming an occluded front. The center of low pressure and of cyclonic rotation is found at the tip of the occluded front, well removed from the warm air source. At this stage the cyclone begins to fill and weaken. *See* STORM.

Many frontal waves are weak, shallow, and fast-moving, appearing more like stage II or III (Fig. 2), rather than progressing to stage IV. Also, cases have been documented in which the occluded structure depicted in panel IV (Fig. 2) forms in a different manner from that described above. In such cases, sometimes referred to as pseudo-occlusions, the low-pressure center is observed to retreat into or form within the cold air, and frontogenesis takes place along a line joining the low center and the tip of the warm sector. Cloud observations from meteorological satellites have provided visual evidence of this process. Since the classical occlusion process,

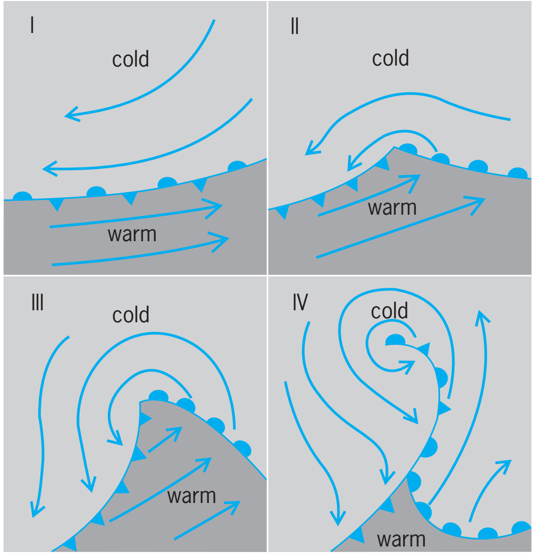


Fig. 2. Life cycle of the wave cyclone, surface projection. Arrows denote airflow. Patterns depicted are for the Northern Hemisphere while their mirror images apply in Southern Hemisphere.

in which the cold front overtakes and merges with the warm front, has never been adequately verified, it is possible that most occlusions form in this other way.

Cloud and precipitation types and patterns bear characteristic relationships to fronts, as depicted in **Fig. 3**. These relationships are determined mainly by the vertical air motions in the vicinity of the frontal surfaces. Since the motions are not unique but vary somewhat from case to case and, in a given case, with the stage of development, the features of the diagram are subject to considerable variation. In general, though, the motions consist of an upgliding of the warm air above the warm frontal surface, a more restricted and pronounced upthrusting of the warm air by the cold front, and an extensive subsidence of the cold air to the rear of the cold front. Fast-moving cold fronts are characterized by narrow cloud and precipitation systems. Smaller, more intense cells or rainbands usually are embedded within the more extensive regions of warm and cold frontal precipitation. *See* CLOUD; CLOUD PHYSICS; PRECIPITATION (METEOROLOGY).

When potentially unstable air is present in the warm sector, as often occurs in the spring and summer over North America, the main weather activity often breaks out ahead of the cold front. Complexes of cumulonimbus clouds, having horizontal dimensions on the order of 60 mi (100 km) and collective lifetimes of perhaps a day, may be present. Convective precipitation may also occur in the form of elongated lines of cumulonimbus clouds called squall lines, as represented by the tall cumulonimbus in Fig. 3*b*. *See* SQUALL; SQUALL LINE.

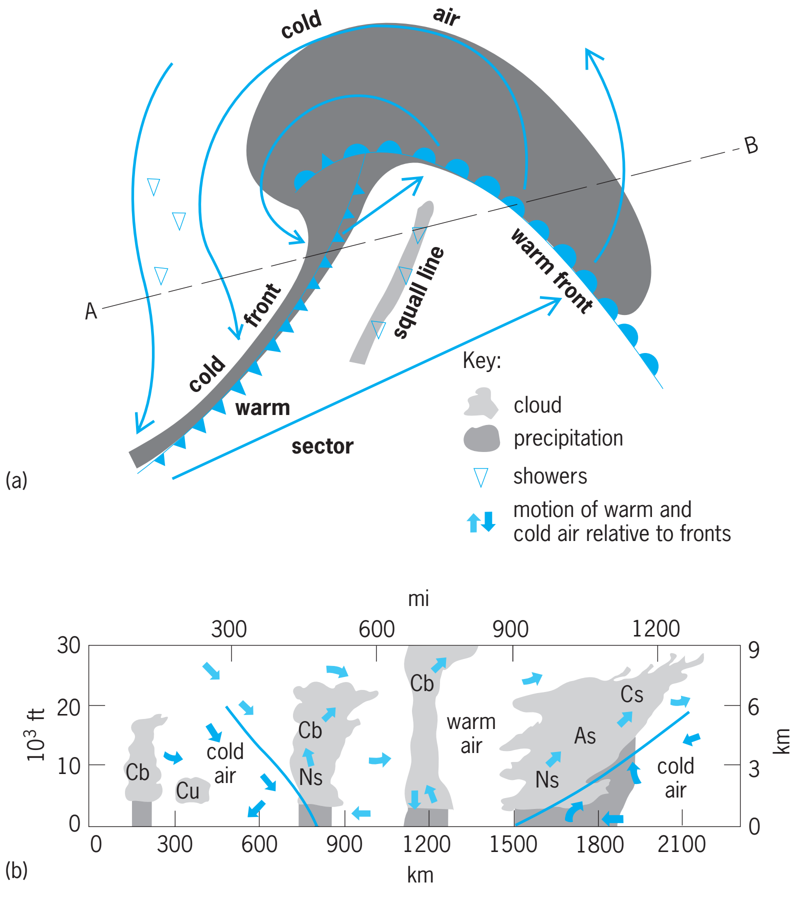


Fig. 3. Relation of cloud types and precipitation to fronts. (*a*) Surface weather map. (*b*) Vertical cross section (along A-B in diagram *a*). Cloud types: Cs, cirrostratus; As, altostratus; Ns, nimbostratus; Cu, cumulus; and Cb, cumulonimbus

In the above conceptual model, the movement of the frontal zones associated with the cyclone's evolution is primarily responsible for the cloudiness and precipitation as the warm air flows up and over the warm front and as the cold front wedges under the warm air sector. Thus, the largely mechanical wedge action of the frontal zones produces the vertical motion associated with the clouds. This conceptual model (known as the polar-front cyclone model) was based on the idea that most cyclones formed on preexisting surface fronts.

Since about 1940, it has been realized that extratropical cyclones often develop in regions where no significant surface frontal zones exist. The development of the cyclones is often more related to the three-dimensional structure of the atmosphere over a large volume than to the location of any frontal zones. More modern conceptual models of cyclone development (which can be referred to as baroclinic-wave models) emphasize the flow of airstreams relative to the evolving cyclone, rather than the distortion of air masses by the moving fronts. Yet the relative distribution of cloudiness and precipitation, as well as the evolution of frontal zones with respect to the cyclone, is still quite similar in these later models to that in the older model discussed above. Thus, Figs. 2 and 3 are consistent with the newer concepts. *See* DYNAMIC INSTABILITY.

Polar front. A front separating air of tropical origin from air of more northerly or polar origin is referred to as a polar front. Frequently only a fraction of the temperature contrast between tropical and polar regions is concentrated within the polar frontal zone, and a second or secondary front appears at higher latitudes. In certain locations such a front is termed an arctic front.

The polar front can also be defined in a climatological sense as a region in which the horizontal temperature gradient tends to be large compared to surrounding areas; within this region frequent formation of sharp frontal zones occurs. This definition applies for the remainder of this section.

In winter the major or polar frontal zones of the Northern Hemisphere extend from the northern Philippines across the Pacific Ocean to the coast of Washington, from the southeastern United States across the Atlantic Ocean to southern England, and from the northern Mediterranean eastward into Asia. An arctic frontal zone is located along the mountain barriers of western Canada and Alaska. In summer the average positions of the polar frontal zones are farther north, the Pacific zone extending from Japan to Washington and the Atlantic zone from New Jersey to the British Isles. In addition to a northward-displaced polar front over Asia, an arctic front lies along the northern shore and continues eastward into Alaska.

The polar frontal zone of the Southern Hemisphere lies near 45°S in summer and slightly poleward of that latitude in winter. Two frontal bands, more pronounced in winter than in summer, spiral into the main zone from subtropical latitudes. These bands, originating east of the Andes and northeast of New Zealand, merge with the main frontal zone after making a quarter circuit of the globe.

Frontogenesis. When extratropical cyclones form in the absence of preexisting fronts, strong frontogenesis (initial formation of a front) can occur in conjunction with the cyclone's development. The formation of a frontal zone requires an increase in the temperature gradient and the development of a wind shift.

The transport of temperature by the horizontal wind field can initiate the frontogenesis process as is shown for two cases in **Fig. 4**. The two wind fields shown would, in the absence of other effects, transport the isotherms in such a way that they would become concentrated along the A-B lines in both cases. (Both of these types of wind fields, or combinations of them, are found in developing extratropical cyclones, and sometimes elsewhere.) Since the temperature gradient is inversely proportional to the spacing of the isotherms, it is clear that frontogenesis would occur along the A-B lines. Vertical air motions will modify this frontogenesis process, but these modifications will be small near the ground where the vertical motion is small. Thus, near the ground the temperature gradients will continue to increase in the frontal zones as long as the horizontal wind field does not change. As the temperature gradient increases, a circulation will develop in the vertical plane through C-D in each case. The thermal wind relation is valid for the component of the horizontal wind which blows parallel to the frontal zone. The thermal wind, which is the change in the geostrophic wind over a specific vertical distance, is directed along the isotherms, and its magnitude is proportional to the temperature gradient. As the frontogenesis process increases the temperature gradient, the thermal wind must also increase. But if the thermal wind increases, the change in the actual wind over a height interval must also increase (since the thermal wind closely approximates the actual wind change with height). This corresponding change in the actual wind component along the front is accomplished through the action of the Coriolis force. A small wind component perpendicular to the front is required if the Coriolis force is to act in this manner. This leads to the circulation in the vertical plane that is shown in **Fig. 5**. *See* CORIOLIS ACCELERATION; CYCLONE; GEOSTROPHIC WIND; THERMAL WIND.

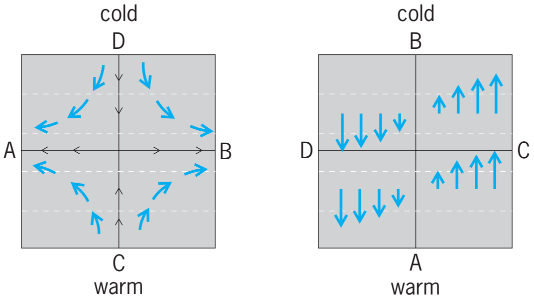


Fig. 4. Two horizontal wind fields which can cause frontogenesis. Broken lines represent isotherms, and colored arrows show wind directions and speeds (length of arrow shaft is proportional to the speed).

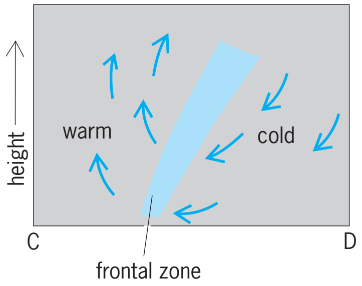


Fig. 5. Circulation in the vertical plane through C-D for both cases of Fig. 4.

This circulation plays an important role in the frontogenesis process and in determining the frontal structure. The rising motion in the warm air and the sinking motion in the cold air are consistent with observed cloud and precipitation patterns. The circulation helps give the front its characteristic vertical tilt which leaves the relatively cooler air beneath the front. Near the ground the circulation causes a horizontal convergence of mass. This speeds up the frontogenesis process by increasing the rate at which the isotherms move together. The convergence field also carries the momentum lines together in the frontal zone in such a way that a wind shear develops across the front. This wind shear gives rise to the wind shift which is observed with a frontal passage. Eventually the front reaches a quasi-steady state in which the turbulent mixing balances the frontogenesis processes. Other frontogenesis effects are important in some cases, but the mechanism presented above appears to be the predominant one.

Upper-level frontal zones. In addition to and independent of the surface frontal zones described above, elongated sloping zones of strong temperature contrast and wind shear can form in the upper troposphere in conjunction with jet stream intensification. These upper-level frontal zones are characterized by low humdity, air sinking within them. Their development is often accompanied by a downward folding of the tropopause and the incorporation of stratospheric air into the upper portion of the zones. The strong wind shear within these frontal zones often produces severe turbulence, a potentially dangerous situation for jet aircraft, as the zones cannot be detected visually due to a lack of cloudiness in the dry, descending air. Upper-level frontal zones thus are significant (but not the only) sources of clear-air turbulence (CAT). *See* CLEAR-AIR TURBULENCE; UPPER-ATMOSPHERE DYNAMICS. **Stephen E. Mudrick**

Bibliography. L. K. Lutgens and E. J. Tarbuck, *The Atmosphere: An Introduction to Meteorology*, 7th ed., 1997; S. E. Mudrick, A numerical study of frontogenesis, *J. Atm. Sci.*, 31:869–892, 1974;

P. Ray (ed.), *Mesoscale Meteorology and Forecasting*, American Meteorological Society, 1987; J. M. Wallace and P. V. Hobbs, *Atmospheric Science: An Introductory Survey*, 1977.

Frost

The formation of ice on the ground, on plants, or on structures near the Earth's surface as a result of the drop in air temperature to lower than 32°F (0°C) and the presence of water vapor or liquid water in the air. The causes of the cooling, and the types of ice that form, can be quite varied, depending on the existing meteorological factors. Frost can have serious impact, with the main effect being in agriculture, but transportation and other areas are also affected. *See* AIR TEMPERATURE.

Causes. There are two principal causes of the cooling necessary for frost formation: heat loss by radiation, and the advection of cold air. Other factors that have a significant influence on frost formation are wind, atmospheric humidity, surface wetness, and topography.

Radiational heat loss. The Earth's surface loses heat by long-wave, infrared radiation. At night, when there is no incoming radiation from the Sun, and especially with clear skies, when there is no returned infrared radiation from clouds, the heat loss can lead to rapid cooling of the surface. In turn, the cold surfaces cool the lowest layers of air as well. In the absence of wind to stir the air, a steep gradient of temperature can become established, so that the temperature within inches of the ground may be many degrees colder than the air several feet above the ground. The depth of this inversion layer increases and the temperature near the surface steadily decreases during the night. *See* TEMPERATURE INVERSION.

Wind tends to cause turbulence and this mixes the coldest air from near the surface with warmer air higher up. Cloudy skies and, to a lesser degree, humidity in the air reduce the rate of heat loss and the risk of frost. Also, conduction of heat from the subsurface soil helps to reduce the loss of heat to some degree. The total amount of cooling that can develop may be calculated by using semiempirical schemes, but the errors in such calculations are fairly large. Furthermore, in hilly terrain the coldest air tends to sink to the lowest elevation, and even small differences in terrain height can lead to substantial differences in temperature. As a result of radiational heat loss, up to 36°F (20°C) cooling during the course of a night is possible.

Heat loss by advection. The movement of cold air, such as that which accompanies the advance of a cold front, is the second major class of situations that can lead to frost formation. In these situations, the air is colder then the ground and, usually, the temperature decreases with altitude. The severity of the drops in temperature can approach those due to radiation cooling, but the most dramatic cooling events are those where both processes contribute. However, fronts are often associated with clouds and usually with winds, so that conditions tend to be contrary to those producing the strongest radiational cooling. *See* FRONT.

Ice types. While low temperatures themselves might be considered as frost, the manifestations and the most serious impacts of frost are in the formation of ice that results from the low temperatures when water is present. The water may be held in the soil or in the plants, or it may come from vapor or droplets in the air.

In order for the water within the surface soil or in the tissues of plants to freeze, a small amount of supercooling is required, typically a few degrees; thus ice formation will not begin (the process is called nucleation) until those temperatures are reached. Nucleation depends on the presence of specific substances. Continued freezing may be limited by the release of latent heat associated with the freezing; in fact some plants have developed protection against extensive freezing damage by having large internal reservoirs of water. Wind velocity is a crucial factor in this regard, as air motion greatly increases the removal of heat from the plant.

Water vapor from the air deposits on surfaces whose temperature is lower than the ice point corresponding to the ambient humidity. These deposits grow as feathery crystals, known as hoarfrost; the crystals are quite beautiful and can attain a size greater than 0.5 in. (1.3 cm). The formation of hoarfrost is analogous to the formation of dew at temperatures higher than 32°F (0°C). *See* DEW; DEW POINT.

If condensation of droplets takes place in the air (that is, a fog forms) coincident with the cooling of the ground and plants, some fog droplets impinge and freeze onto the objects, forming rime. Rime is much denser than hoarfrost. The total mass (weight) deposited by rime can also be much higher than the deposit of hoarfrost. Occasionally, the two forms of ice form in combination. *See* FOG.

If moisture is deposited as liquid water prior to the lowering of temperatures below the freezing point, the water deposit (dew or rain) may subsequently freeze, forming a transparent (black) ice coating known as glaze. Very heavy and damaging glaze can form if rain falls from warmer air aloft into a surface layer of cold air. *See* PRECIPITATION (METEOROLOGY).

Climatology. Frost climatology is widely available for agricultural areas. The dates of first and last frosts for 32, 28, 24, 20°F (0, −2.2, −4.4, −6.7°C), or other sets of temperatures, are usually tabulated, as well as the probabilities of frosts for given calendar days. Tables of the numbers of frost days per year are also available. These data represent the average conditions over many years. Deviations from the average can always be expected. Also, local variations are very important. Other factors, such as humidity or water content in the soil, add further variability to the actual frequency and impact of frosts.

Forecasting and prevention. General advances in the 1- to 2-day synoptic-scale meteorological forecasts are providing improved information for the generation of frost forecasts. As a result, frost warnings of several hours are quite reliable, but forecasts more than a day ahead are still beset with inaccuracies. The incorporation of small-scale local effects places further demands on the forecaster. Good forecasts of frost are of great value to growers who operate expensive frost-protection systems. Prevention measures are decided on the basis of the forecasts and the susceptibility of the plants at that moment. Common techniques of frost prevention are heating of the air with burners, smoke production to reduce radiation heat loss, spraying of the plants with water to utilize the latent heat of freezing as a heat source, and treatments to reduce the temperature threshold for nucleation. The last-mentioned method is experimental, but it has the potential of giving protection down to 24°F (−4.4°C). *See* AGRICULTURAL METEOROLOGY.

Gabor Vali

Froude number

The dimensionless quantity $U(gL)^{-1/2}$, where U is a characteristic velocity of flow, g is the acceleration of gravity, and L is a characteristic length. The Froude number can be interpreted as the ratio of the inertial to gravity forces in the flow. This ratio may also be interpreted physically as the ratio between the mean flow velocity and the speed of an elementary gravity (surface or disturbance) wave traveling over the water surface. *See* WAVE MOTION IN LIQUIDS.

In an open channel, the characteristic length is taken to be the hydraulic depth, which by definition is the cross-sectional flow area divided by the width of the free surface. For rectangular sections, the characteristic length equals the flow depth. *See* OPEN CHANNEL.

When the Froude number is equal to one, the speed of the surface wave and that of the flow is the same. The flow is in the critical state. When the Froude number is less than one, the flow velocity is smaller than the speed of a disturbance wave traveling on the surface. Flow is considered to be subcritical (tranquil flow). Gravitational forces are dominant. The surface wave will propagate upstream and, therefore, flow profiles are calculated in the upstream direction. When the Froude number is greater than one, the flow is supercritical (rapid flow) and inertial forces are dominant. The surface wave will not propagate upstream, and flow profiles are calculated in the downstream direction.

The dimensionless wave drag and power required for a surface ship are strong functions of its Froude number. Therefore, tow-tank models are tested at the same Froude number as the prototype ship. The Reynolds number of the model does not match the prototype, and viscous drag is generally estimated in tow-tank tests. *See* REYNOLDS NUMBER; SHIP POWERING, MANEUVERING, AND SEAKEEPING; TOWING TANK.

The Froude number is useful in calculations of hydraulic jump, design of hydraulic structures, and ship design, where forces due to gravity and inertial forces are governing. In these cases, geometric similitude and the same value of the Froude number in model and prototype produce a good approximation to dynamic similitude. *See* DIMENSIONAL ANALYSIS; DIMENSIONLESS GROUPS; HYDRAULICS; SHIP DESIGN.

Raymond M. Wright

Bibliography. N. H. C. Hwang and R. J. Houghtalen, *Fundamentals of Hydraulic Engineering Systems*, 3d ed., Prentice Hall, 1996; V. L. Streeter, E. B. Wylie, and K. W. Bedford, *Fluid Mechanics*, 9th ed., McGraw-Hill, 1998; T. Szirtes and P. Rozsa, *Applied Dimensional Analysis and Modeling*, McGraw-Hill, 1997; R. B. Zubaly, *Applied Naval Architecture*, Cornell Maritime Press, 1996.

Fruit

A matured carpel or group of carpels (the basic units of the gynoecium or female part of the flower) with or without seeds, and with or without other floral or shoot parts (accessory structures) united to the carpel or carpels. Carpology is the study of the morphology and anatomy of fruits. The ovary develops into a fruit after fertilization and usually contains one or more seeds, which have developed from the fertilized ovules. Parthenocarpic fruits usually lack seeds. Fruitlets are the small fruits or subunits of aggregate or multiple fruits. Flowers, carpels, ovaries, and fruits

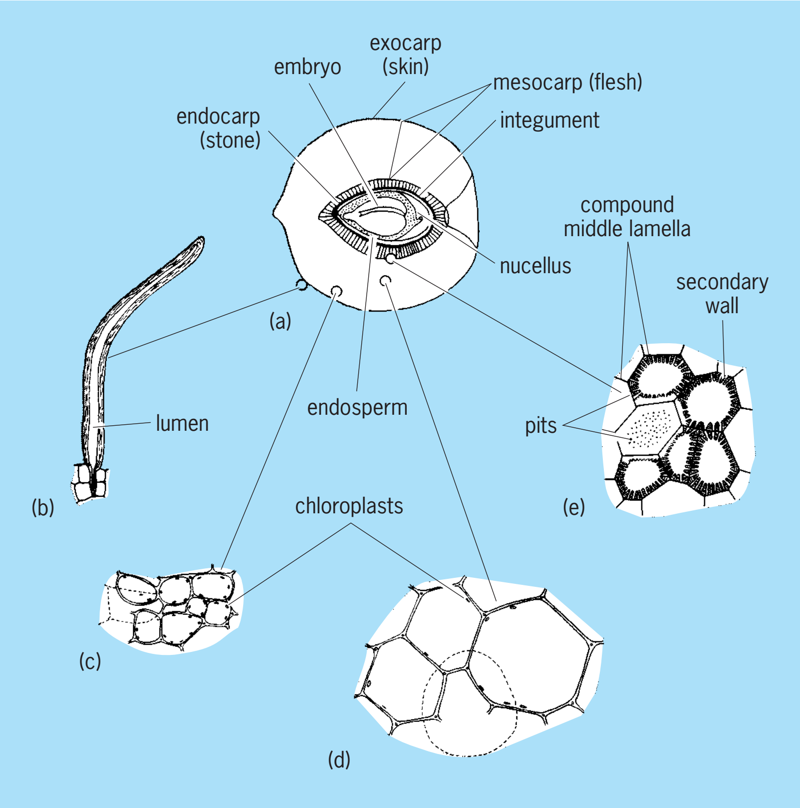


Fig. 1. Tissue components of fruit of *Prunus persica* (peach). (*a*) Longitudinal section of fruit, a drupe. Circles indicate positions of sections shown in *b–e*, which are representative of a young fruit about 1 in. (2.5 cm) in diameter. (*b*) Epidermal hair of exocarp. (*c*, *d*) Parenchyma of fleshy mesocarp. (e) Sclereids of stony endocarp. (*After K. Esau, Plant Anatomy*, *2d ed*., *John Wiley and Sons, 1967*)

are, by definition, restricted to the flowering plants (angiosperms), although fruitlike structures may enclose seeds in certain other groups of seed plants. The fruit is of ecological significance because of seed dispersal. *See* FLOWER; SEED.

Morphology. A fruit develops from one or more carpels. Usually only part of the gynoecium, the ovary, develops into a fruit; the style and stigma wither. Accessory (extracarpellary or noncarpellary) structures may be closely associated with the carpel or carpels and display various degrees of adnation (fusion) to them, thus becoming part of the fruit. Such accessory parts include sepals (as in the mulberry), the bases of sepals, petals, and stamens united into a floral tube (apple, banana, pear, and other species with inferior ovaries), the receptacle (strawberry), the pedicel and receptacle (cashew), the peduncle (fleshy part of the fig), the involucre composed of bracts and bracteoles (walnut and pineapple), and the inflorescence axis (pineapple).

True versus accessory fruits. A fruit derived from only carpellary structures is called a true fruit or, because it develops from a superior ovary (one inserted above the other floral parts), a superior fruit (**Fig. 1**; corn, date, grape, plum, and tomato). Fruits with accessory structures are called accessory (or inaptly, false or spurious) fruits (pseudocarps) or, because of their frequent derivation from inferior ovaries (inserted below the other floral parts), inferior fruits (**Fig. 2**; banana, pear, squash, and walnut). True fruits invariably develop from the superior ovaries of hypogynous or perigynous flowers; accessory fruits usually develop from the inferior or part-inferior ovaries of epigynous or semiepigynous flowers. Exceptions include the accessory fruits of the rose and strawberry, which contain achenes (true fruitlets) that develop from the carpels of perigynous flowers. The fruit of the rose (the so-called hip) represents a floral tube enclosing the achenes. A strawberry is a large, fleshy receptacle bearing small achenes. True and accessory fruits are distinguishable on the basis of developmental anatomy.

Simple and aggregate versus multiple fruits. Fruits can be characterized by the number of ovaries and flowers forming the fruit. A simple fruit is derived from one ovary (Figs. 1–5), an aggregate fruit from several ovaries of one flower (**Fig. 3***a* and *b*; magnolia, rose, and strawberry). A multiple (collective) fruit is derived from the ovaries and accessory structures of several flowers consolidated into one mass (Fig. 3*c*;

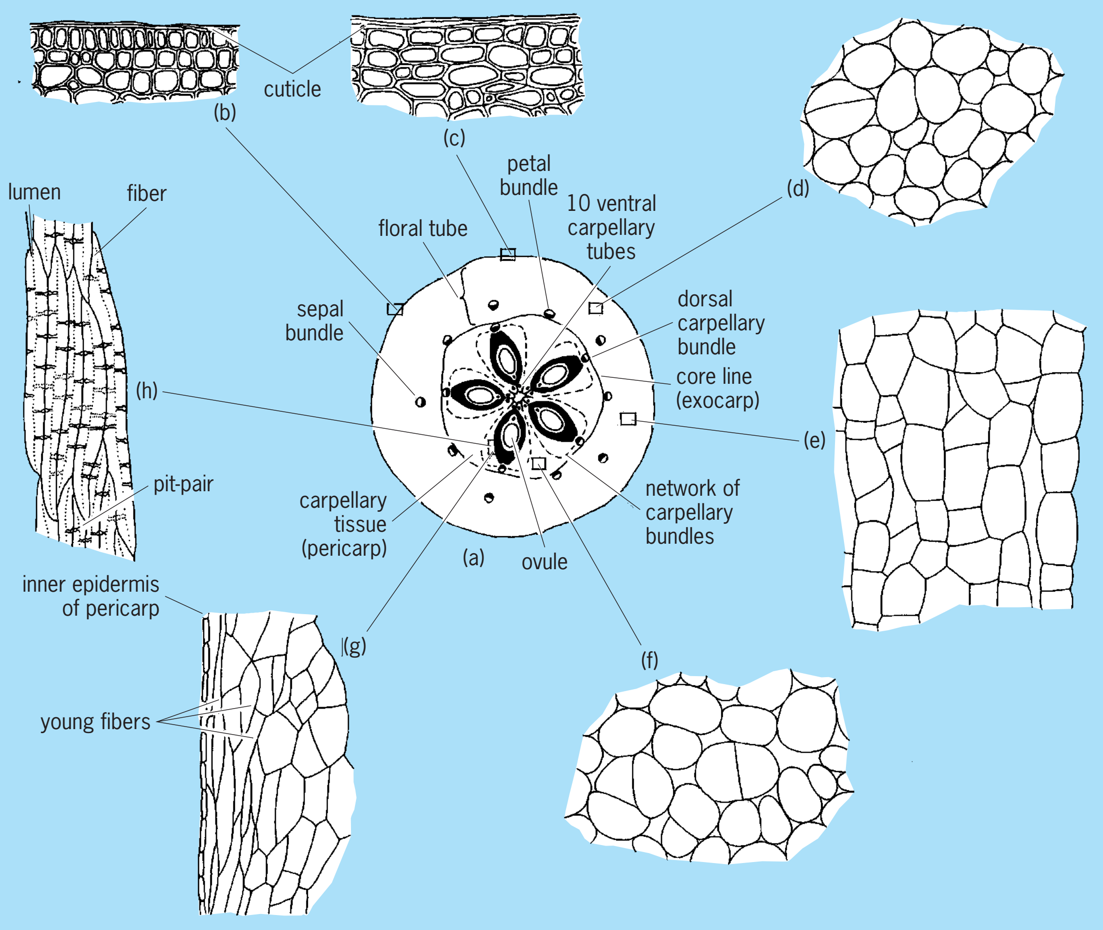


Fig. 2. Tissue components of fruit of *Malus sylvestris* (apple). (*a*) Transverse section of fruit, a pome. Rectangles indicate positions of sections shown in *b–h*; *b* and *d–g* are from a young fruit 0.4 in. (1 cm) in diameter; *c* and *h* are from a mature fruit. (*b*) Epidermis and subjacent collenchyma from young fruit and (*c*) from mature fruit. (*d, e*) Parenchyma of floral-tube part of flesh. (*f*) Parenchyma of mesocarp part of flesh. (*g*) Fibers of cartilaginous endocarp from young fruit and (*h*) from mature fruit. (*After K. Esau, Plant Anatomy*, *2d ed*., *John Wiley and Sons*, *1967*)

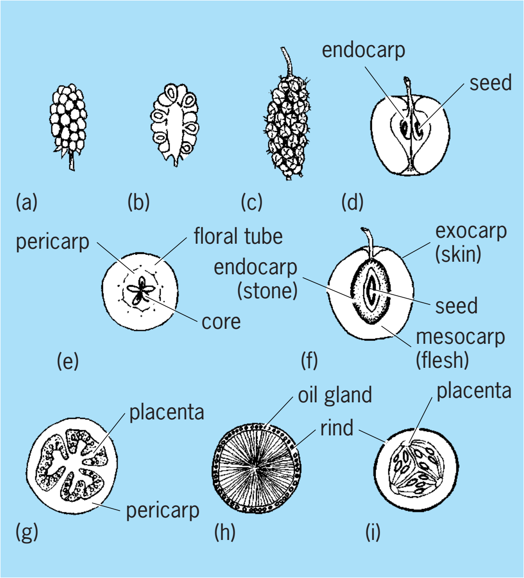


Fig. 3. Fleshy fruits. (*a*) Aggregate fruit (blackberry) in surficial view and (*b*) in sectional view. (*c*) Multiple fruit (mulberry). (*d–i*) Simple fruits: (*d*, *e*) pome (apple), (*f*) drupe (cherry), (*g*) berry (tomato), (*h*) hesperidium (orange), (*i*) pepo (cucumber); *d* and *f* are longitudinal sections, *e* and *g–i* are transverse sections.

fig, pandan, pineapple, and sweet gum). Simple fruits are either true (Fig. 1) or accessory (Fig. 2) fruits. However, aggregate and multiple fruits are always accessory fruits. Their individual units or fruitlets are characterized by the terminology of simple fruits given below. In an aggregate fruit a fruitlet is the individual carpel of a flower with separate carpels, whereas in a multiple fruit a fruitlet is the ovary and accessory structures of a flower. Simple, aggregate, an multiple fruits may be fleshy or dry, and dehiscent or indehiscent.

Dry versus fleshy, and dehiscent versus indehiscent fruits. The fruit wall at maturity may be fleshy or, more commonly, dry. Fleshy fruits range from soft and juicy to hard and rough. Dry fruits may be dehiscent, opening to release seeds, or indehiscent, remaining closed and containing usually one seed per fruit. Dehiscent fruits have areas of opening called sutures and separating segments called valves (**Fig. 4**). Fleshy fruits are rarely dehiscent.

Fruit wall versus pericarp. The pericarp is the fruit wall developed from the ovary. In true fruits, the fruit wall and pericarp are synonymous (Fig. 1), but in accessory fruits the fruit wall includes the pericarp plus one or more accessory tissues of various derivation (Fig. 2). Some persons inappropriately broaden the term pericarp to include noncarpellary tissue. The pericarp is usually divisible anatomically into the exocarp, mesocarp, and endocarp. Besides the fruit wall, a fruit contains one or more seed-bearing regions (placentae) and often partitions (septa).

Classification. Fruits have been classified on the basis of numerous criteria, including dehiscence versus indehiscence, dry versus fleshy texture, morphology (form), development (ontogeny), relationship of the ovary to other reproductive parts, and number of carpels and seeds. The terminology and classification presented here is the traditional one developed by taxonomists and temperate botanists and is biased toward fruits of economic importance. Although these fruit types are widely accepted, the classification is artificial since numerous transitional types occur and since morphological, histological, developmental, ecological, and functional criteria are used rather indiscriminately. Nevertheless, the numerous attempts at developing a natural system of fruit classification have not been very successful.

Dehiscent dry fruits. Dehiscent fruits, sometimes uncritically called pods, usually have a dry texture at maturity. The following types (Fig. 4) are representative: follicle, containing one carpel, opening along one suture (milkweed and fruitlet of magnolia); legume, with one carpel, but opening along two sutures (common garden bean and many other members of the legume family); loment, a special type of legume constricted between seeds and breaking crosswise into one-seeded segments (beggar's-ticks); silique, composed of two elongated carpels with two sutures, the valves opening from below upward to expose the seed-bearing partition (replum) still attached to the stalk (many members of the mustard family); silicle, like a silique, but short and broad (shepherd's-purse); schizocarp, splitting into one-seeded usually indehiscent units (mericarps or sometimes "cocci"), each a derivative of a carpel—sometimes regarded as intermediate between dehiscent and indehiscent dry fruits (mallow and members of the unbel family); capsule, containing two or more carpels, and usually with a like number of vertical valves or sutures (jojoba and lily). Capsules have diverse modes of dehiscence. Thus, in the pyxis subtype dehiscence is circumferential, with the top separating as a lid (purslane).

Indehiscent dry fruits. Representative types (**Fig. 5**) include caryopsis or grain, with the fruit wall fused to the single seed (corn and other members of the grass family); achene (sometimes akene), with the fruit wall readily separable from the single seed (clematis and the fruitlets of a strawberry; cypselae are

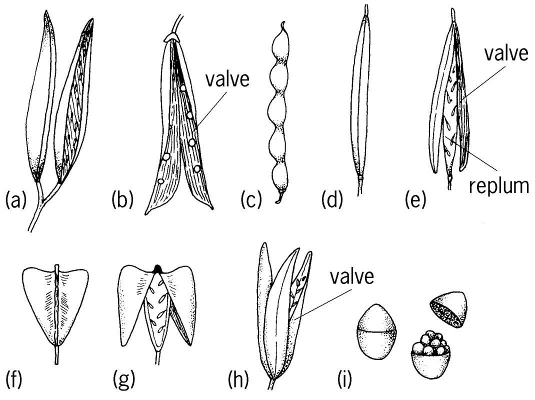


Fig. 4. Dehiscent dry fruits: (*a*) Follicles (milkweed). (*b*) Legume (bean). (*c*) Loment (beggar's-ticks). (*d*, *e*) Silique, closed and open (mustard). (*f*, *g*) Silicle, closed and open (shepherd's-purse). (*h*) Capsule (lily). (*i*) Pyxis, closed and open (purslane).

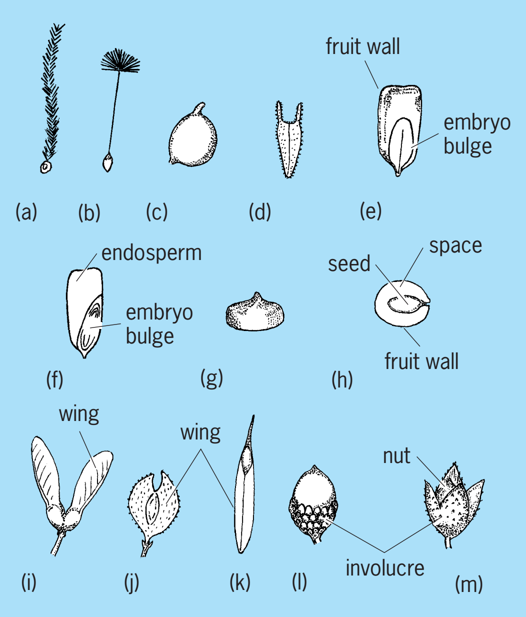


Fig. 5. Indehiscent dry fruits. (*a–d*) Achenes; (*a*) clematis, (*b*) dandelion, (*c*) buttercup, and (*d*) sticktight. (*e*) Caryopsis or grain (corn) in surficial view and (*f*) in sectional view. (*g*) Utricle (goosefoot) in surficial view and (*h*) in sectional view. (*j–k*) Samaras: (*i*) maple, (*j*) elm, and (*k*) ash. (*l*) Nut and involucre (acorn) of oak. (*m*) Nut of beech.

achenes derived from inferior ovaries, as in members of the sunflower family); utricle, like an achene, but with the fruit wall thin and bladdery (beet and goosefoot); samara, like an achene, but with a wing for wind dispersal (ash, elm, and maple); nut, like an achene but larger, with the fruit wall hard throughout, derived from several carpels, and usually one-seeded and often at least partly enclosed in an involucre (beech, hazelnut, walnut, and the acorn of oak).

Fleshy fruits. Fleshy fruits are usually indehiscent, but there are exceptions, such as squirting cucumber. The chief types (Fig. 3) of simple fleshy fruits include pome, with a thin, cartilaginous inner part (endocarp) and an fleshy outer part (Fig. 2; pear and pomegranate, the latter a leathery pome or balausta); drupe (stone fruit), with an outer exocarp (skin) in the fruit wall, a central fleshy mesocarp (flesh), and an inner stony endocarp or stone (Fig. 1; cherry); berry, with the entire fruit fleshy (avocado, banana, grape, kiwi, and tomato); hesperidium, a modified berry with a thick, leathery separable rind (citrus fruits); pepo, a modified berry with a hard or tough inseparable rind (cucumber and other members of the gourd family).

Anatomy. Anatomically or histologically, a fruit consists of dermal, ground (fundamental), and vascular systems and, if present, one or more seeds. After fertilization the ovary and sometimes accessory parts develop into the fruit; parthenocarpy is fruit production without fertilization. The fruit generally increases in size and undergoes various anatomical changes that usually relate to its manner of dehiscence, its mode of dispersal, or protection of its seeds. The economically important, mainly fleshy fruits have received the most histological and developmental study. *See* COLLENCHYMA; EPIDERMIS (PLANT); PARENCHYMA; SCLERENCHYMA.

Fruit growth. Size increase of fruits is hormonally controlled and results from cell division and especially from cell enlargement. Cell number, volume, and weight thus control fruit weight. Cell division generally is more pronounced before anthesis (full bloom); cell enlargement is more pronounced after. Fruit growth may result only from cell enlargement, all cells of the mature fruit being present at anthesis (corinth grape and raspberry). More often the fruit develops by cell enlargement accompanied by cell division, which may continue briefly (tomato, and peach and other stone fruits) or until maturity (avocado and citrus). Cells may enlarge 10–15 times (kiwi) or more. Expansion of intercellular spaces may also contribute substantially to fruit growth, for instance, 25% of the total volume in apples. The growth curve may be a single sigmoid (apple, pea, and tomato) or double sigmoid (grape, kiwi, and stone fruits). Usually not all parts of the fruit grow at the same rate. Cell layers present at anthesis or during early fruit development may later become obscured.

Fruit wall and pericarp. As defined above, the pericarp or fruit wall of true fruits represents the ovary wall, whereas the fruit wall of accessory fruits includes the pericarp plus accessory noncarpellary tissue. The pericarp is subdivided into the exocarp, mesocarp, and endocarp, but these designations are purely descriptive and have no real developmental basis. The pericarp layers may merge imperceptibly into one another, and only endocarp and exocarp may be present. In accessory fruits the pericarp may be anatomically indistinguishable from noncarpellary tissue (lettuce). However, in some apple cultivars (pink pearl) these regions are distinguishable: the pericarp is white, and the accessory tissues are pink. Anatomically, fruit walls of indehiscent dry fruits often resemble seed coats (in the achene of sunflower and caryopses of grasses). Structures other than the fruit wall may constitute a principal part of a fruit, for example, expanded placentae (Fig. 3*g*; tomato) or fleshy seed coats (pomegranate). True fruits may be slightly (kiwi) or appreciably (green ash) surrounded by a persistent calyx.

Epidermis and parenchyma are always present, and vascular tissue nearly always present in the fruit wall (fruits of *Balanophora* completely lack vascular tissue), whereas collenchyma and especially sclerenchyma—fibers or sclereids or both—are often present, sometimes forming rather late in fruit development (**Figs.** 1, 2, **6**, and **7**). The flesh in mature fleshy fruits is due to succulent parenchyma cells, which contain water and solutes mainly in their vacuoles. Immature dry fruits also exhibit some fleshiness. Parenchyma cells may be tightly packed (Figs. 1 and 6) or loosely arranged with small or often large intercellular spaces (Fig. 2). Mealiness or mushiness of overripe fruits results from pronounced cell separation (apple). Complete cell separation occurs in the flesh of some fruits (cantaloupe). Collenchyma

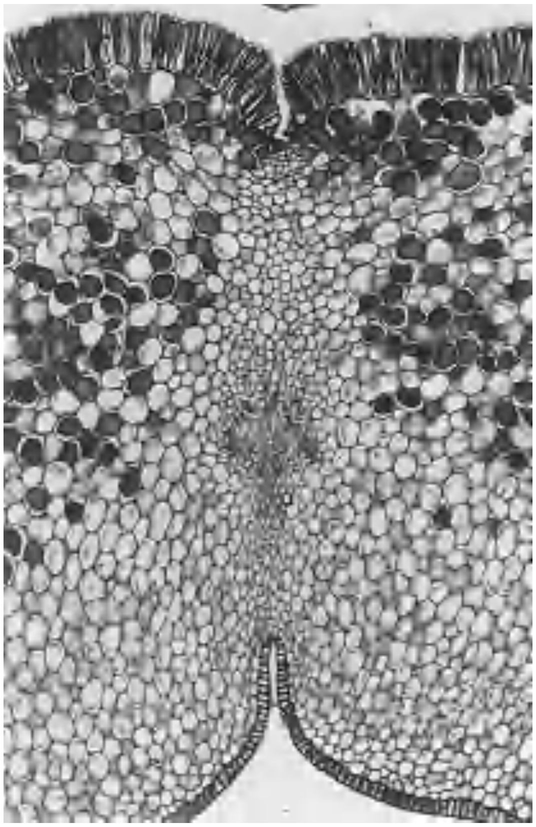


Fig. 6. Transverse section of wall of immature fruit of *Simmondsia chinensis* (jojoba) at point of potential dehiscence (central smaller cells). The outer epidermis (at the top) consists of tannin-filled macrosclereids. (*From R. Schmid, Floral and fruit anatomy of jojoba, Simmondsia chinensis, Mem. II. Conf. Int. Jojoba Aprovech., Ensenada, Mexico, pp. 143–148, 1978*)

and sclerenchyma occur in both dry and fleshy fruits, though usually in greater abundance in dry fruits. Pear flesh represents noncarpellary tissue consisting of parenchyma with vascular tissue and islands of sclereids. Dry fruits are generally more complex than fleshy ones, in part because many dry fruits dehisce whereas fleshy fruits rarely do. Tannins (Fig. 6), crystals, starch grains, latex, oils, fats, proteins, chloroplasts, pigments (anthocyanins and carotenoids) regularly occur in many fruit walls. Both true fruits (kiwi and sapote) and accessory fruits (apple and pear) may be russeted from cork development in some of the outer cell layers. Functional nectaries (trumpet creeper) and hydathodes sometimes occur on fruits.

Exocarp. In accessory fruits the exocarp (sometimes called the epicarp) may merge imperceptibly with the noncarpellary tissue of the other fruit wall, although some multivars of apple exhibit an exocarp or core line (Fig. 2) of smaller cells. In true fruits the exocarp may consist only of epidermis (Fig. 6; pea) or may be multilayered and thus thick and distinct (Fig. 1; the skin of plum). A thin or thick cuticle is invariably present. The exocarp may be smooth (Fig. 6; avocado and grape) or pubescent, with unicellular or multicellular hairs, or both (Fig. 1; apricot and kiwi). Stomata may be present (citrus and kiwi) or absent (grape). In mature fruits lenticels may replace stomata (horse chestnut). The epidermal cells may develop into macrosclereids (Fig. 6). Often beneath the epidermis is a hypodermis or a subepidermal region (or both) one to several layers thick and containing pigments, tannins, crystals, starch grains, oil cells, oil cavities, collenchyma, fibers, or sclereids. Vascular tissue is usually absent. The exocarp may be fleshy or dry, and its cells may have thin or thick walls. The outer layers of noncarpellary tissue of accessory fruits may have the same structure as exocarps of true fruits. As noted above, russeting sometimes results from cork formation.

Mesocarp. The mesocarp (sometimes called the sarcocarp) occurs between the exocarp and endocarp. Other than the vascular tissue that is typically present, the mesocarp may contain only homogeneous parenchyma (Figs. 1, 2, and 6; grape and kiwi),

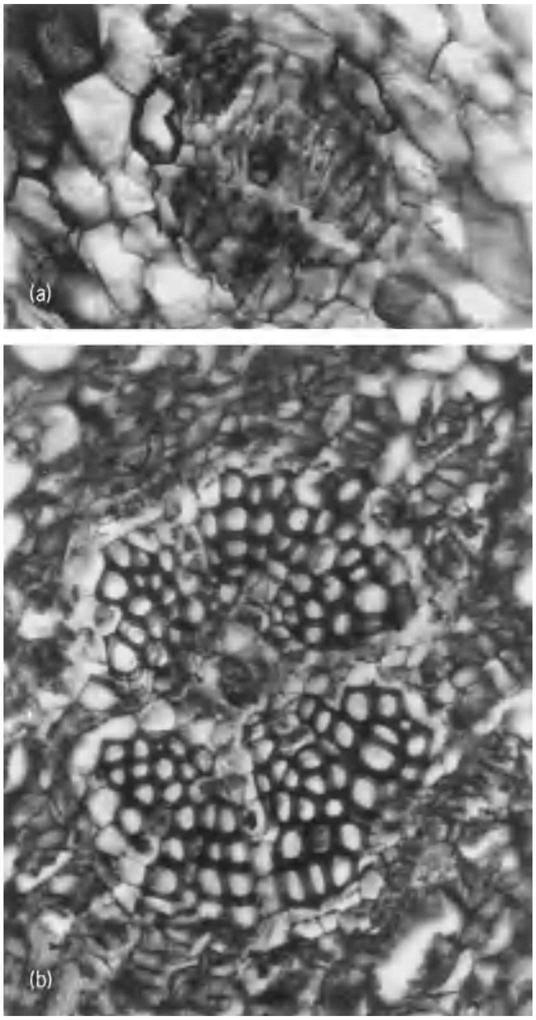


Fig. 7. Transverse sections of placental vascular bundles of (*a*) female flower bud and (*b*) mature fruit of *Actinidia chinensis* (kiwi). Both bundles are from comparable parts of the flower or fruit, both are mesarch amphicribral, and both are at the same magnification. The fruit bundle in *b* has a parenchymatous center and apparent secondary growth. During fruit development the large ovary bundles increase in size 40–80%. (*From R. Schmid, Reproductive anatomy of Actinidia chinensis (Actinidiaceae), Bot. Jahrb. Syst., 100:149–195, 1978*)

or intermixed parenchyma and sclerenchyma, especially fibers (legumes), or intermixed parenchyma and collenchyma (legumes). With maturity some parenchyma may become disorganized and show less wall thickening (peach and tomato). Sometimes mesocarp and endocarp are indistinguishable from each other.

Endocarp. The endocarp may be simply a thin- or thick-walled epidermis next to the ovary locule and seed (Fig. 6), but usually it is more complex (Figs. 1 and 2). The endocarp may be parenchymatous and anatomically like the mesocarp, but usually the endocarp is distinct, as in the highly sclerenchymatous stone and pome fruits (Figs. 1 and 2), or in legumes with layers of fibers often oriented crosswise. Collenchyma is sometimes present (legumes), and vascular tissue is frequently so. The epidermis often becomes sclerenchymatous and may bear a cuticle, hairs, stomata, or, in citrus fruits, outgrowths that are the juice vesicles.

Vascular system. The fruit wall, especially of large fruits, often contains much vascular tissue, either few large bundles or many small ones, or both. Frequently additional vascular bundles are differentiated to supplement the basic vascular pattern of the flower. Vascular tissue may occur throughout the pericarp and, in accessory fruits, also in noncarpellary tissue (Fig. 2). Often vascular bundles anastomose to form a complex network in the fruit wall (Fig. 2; kiwi). Fruit vascular bundles commonly are mesarch and amphicribral, with phloem surrounding xylem (Fig. 7), or endarch and collateral, with increased amounts of phloem (Fig. 2). Some bundles may be strictly phloic, whereas others may increase in diameter during fruit development by primary or secondary growth, or both (Fig. 7). *See* PHLOEM; XYLEM.

Abscission. Unless anatomically united with the carpel or carpels, floral parts usually abscise (separate and fall) during fruit development. Thus the ovary frequently is the only floral structure remaining at maturity. Sometimes perianth parts, especially sepals, are persistent. Fruits usually abscise from the plant when ripe and are dispersed (coconut, peach, and sycamore). The fruit and its stalk (peduncle or pedicel) may separate as a unit (apple and pear), or the stalk may remain on the branch when the fruit abscises (citrus and peach). The abscission zone, if present, may involve either the formation of a definite layer of corklike cells across the stalk or at the base of the fruit, or cytological changes in existing cells, especially in and between their cell walls. Some cells show no changes at abscission but are easily broken mechanically. Some fruits abscise without a definable abscission zone. *See* ABSCISSION.

Functional aspects. Most fruits increase in size compared to the flower, sometimes dramatically so (300,000 times in avocado). Large fruits generally require additional anatomical modifications for nutrition or support or both. The extra phloem in fruit vascular bundles and the often increased amount of vascular tissue in the fruit wall and septa supply nutrients to the developing seeds and, especially in fleshy fruits, to the developing walls. Large, especially fleshy fruits (apple, gourd, and kiwi) usually contain proportionally more vascular tissue than small fruits (achenes and follicles). Vascular tissue also serves for support and in lightweight fruits may be the chief means of support. Collenchyma and particularly sclerenchyma are common supportive and protective tissues. Often the long axes of cells are oriented crosswise to give a mechanically strong structure.

Crystals, tannins, and oils commonly occur in fruits and may protect against pathogens and predators. The astringency of tannins, for example, may be repellent to organisms. With fruit maturation, tannin content ordinarily decreases, so the tannin repellency operative in early stages is superseded in fleshy fruits by features (tenderness, succulence, sweetness through odor and increased sugar content, and so on) attractive to animal dispersal agents. Many fruits are dispersed by hairs, hooks, barbs, spines, and sticky mucilage adhering the fruit to the surface of the dispersal agent. Lightweight fruits with many air spaces or with wings or plumes may be dispersed by wind or water. Gravity is always a factor in dispersal of fruits and seeds. Dehiscent fruits release seeds by varied and often complicated mechanisms. For example: increased turgor pressure may eventually cause the fruit to explode (squirting cucumber and touch-me-not); water loss by crossed superposed layers of cells, especially fibers, may set up antagonistic forces eventually causing the fruit to open (legumes), or ordinary cell enlargement may effect the same (jojoba). Small thin-walled cells often mark where the fruit wall actually splits (Fig. 6). *See* PLANT ANATOMY; REPRODUCTION (PLANT).

Rudolf Schmid

Bibliography. B. G. Coombe, The development of fleshy fruits, *Annu. Rev. Plant Physiol.,* 27:207–228, 1976; A. Fahn and E. Werker, Anatomical mechanisms of seed dispersal, in T. T. Kozlowski (ed.), *Seed Biology,* vol. 1, 1972; I. Roth, Fruits of angiosperms, in Linsbauer's *Handbuch der Pflanzenanatomie*, vol. 10(1), 2d ed., 1977; R. Schmid, Functional and ecological interpretations of floral and fruit anatomy, *Bot. Rev.*, vol. 47, 1981.

Fruit, tree

Tree fruits include temperate, subtropical and tropical zone species. Most temperate zone fruits are deciduous, that is, they lose their leaves in the autumn. They are grown principally in regions protected from prolonged summer heat and severe winter cold (above −10 to −15°F or −23 to −26°C). The principal deciduous tree fruits grown in the United States are apple, peach, pear, plum, apricot, sweet cherry, tart cherry, and nectarine. Tree nuts, such as almond, pecan, walnut, and filbert, are sometimes classified as deciduous tree fruit crops.

Most subtropical fruit trees are evergreen. They will withstand temperatures somewhat below freezing during their dormant or semidormant season, but not the extreme temperatures tolerated by the temperate zone crops. Major subtropical fruits in the

United States are the citrus group (for example, orange, grapefruit, lemon), olive, avocado, fig, and others of lesser importance. The Japanese persimmon might be considered as a borderline case between the temperate and subtropical zone groups, and avocado is often classed as a tropical fruit.

Tropical fruits are killed or severely injured by temperatures below freezing and some are sensitive to cold temperatures above freezing. For example, papayas will survive in the milder areas of the subtropics but fail to mature fruits of good quality. The principal tropical fruits grown in the continental United States are date and mango, while banana and papaya are important in its island possessions. See separate articles on the common-name fruits and nuts.

Areas of production. California produces over 50% of the total United States deciduous fruit tonnage and is the number one state in production of peaches, pears, plums, apricots, and nectarines, as well as an important producer of apples and sweet cherries. Washington is first in production of apples and sweet cherries, while Michigan ranks first in tart cherry production. California also ranks first in production of lemons, olives, avocados, figs, dates, almonds, walnuts, grapes, and several other orchard crops. Florida is first in production of oranges and grapefruits.

Propagation. Nearly all tree fruits, whether deciduous or tropical, are propagated asexually by budding or grafting the desired variety onto the selected rootstock. The rootstock is usually a young seedling or a clonal stock. *See* PLANT PROPAGATION.

Most deciduous tree fruits are grown on seedlings of selected species. Trees grown on seedling stocks attain normal size, as opposed to dwarfing or semidwarfing stocks which have gained favor, particularly with the apple. Seedlings of peach, apricot, and several plum species (such a mariana and myrobalan) are used as stocks for peach, apricot, and plum. Mazzard and mahaleb cherry seedlings are used for most sweet and tart cherries. Apple and pear seedlings are used for standard-sized trees of these two fruits, while selected dwarfing clones of apple are used for dwarf apple trees, and quince is used as a pear dwarfing stock. Several citrus species are used as rootstocks for the various citrus fruits.

Cultivation. The tree fruit crops are grown commercially in orchards or groves, usually in single rows which permit necessary cultural operations for each tree. Nearly all tree fruit crops have similar requirements of training, pruning, spraying to control diseases and insects, and cultivation or chemical control of weeds. Most tree fruits receive applications of nitrogen and other nutrients as required. In areas of low rainfall, orchards are irrigated.

Most orchard crops are harvested by hand, although mechanical harvesting has been successful for some processing crops, including tart cherries and plums for drying (prunes). *See* AGRICULTURAL MACHINERY.

R. Paul Larsen

Diseases. Tree fruits have diseases that may spoil the appearance of the fruits or render them unfit for human consumption; diseases may also lower yield or shorten the productive life of the trees. Most of the tree-fruit diseases are highly specialized, attacking specific hosts. Only infrequently will one tree-fruit disease attack a wide range of fruits.

Fig. 1. Apple scab on young fruit.

Temperate-zone fruits. Diseases of temperate-zone fruits (sometimes called deciduous fruits) are due to fungi, bacteria, viruses, and adverse environmental conditions.

Fungi are the most frequent causes of temperate-zone tree-fruit diseases. For example, a perennial threat to most commercial and home apple orchards is the apple scab disease (**Fig. 1**). Caused by the fungus *Venturia inaequalis*, scab may cause premature defoliation, June drop of young fruits, and unsightly blemishes on ripe apples. A similar disease attacks most varieties of pears and is caused by *V. pyrina*. *See* FUNGI.

Stone fruits, such as peaches, plums, cherries, apricots, and nectarines, are frequently damaged extensively by brown rot caused by the fungus *Monilinia fructicola* (**Fig. 2**). This disease attacks the flowers (causing blossom blight), infects terminals and twigs with twig blight, and is widely recognized by its destruction of ripe fruit at harvest time.

Many diseases caused by fungi attack and destroy leaves of tree fruits: cherry leaf spot, *Coccomyces hiemalis;* peach leaf curl, *Taphrina deformans;* apple cedar rust, *Gymnosporangium juniperi-virginianae;* leaf blight of pears, *Mycosphaerella sentina;* and apple powdery mildew, *Podosphaera leucotricha*.

Bacteria also attack tree fruits, causing extensive damage in some regions. For example, fireblight of apples and pears, *Erwinia amylovora*, is a common disease of pomaceous fruits in most seasons (**Fig. 3**). Bacterial leaf spot of stone fruits, *Xanthomonas pruni*, is often the cause of serious damage to leaves,

Fig. 2. An example of brown rot on plums.

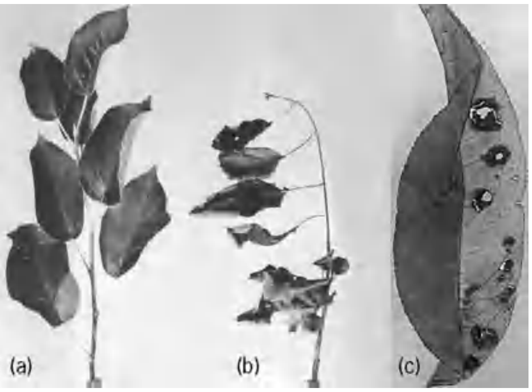


Fig. 3. Several examples of bacterial plant diseases. (*a*) Healthy apple twig and (*b*) fireblight-diseased apple twig. (*c*) Bacterial leaf spot of peach.

twigs, and fruits of peaches, apricots, almonds, cherries, plums, and nectarines. Crown gall, *Agrobacterium tumefaciens*, may infect tree fruits in the nursery row, causing the formation of large galls on roots and stems of young apple, peach, and pear trees. *See* BACTERIA.

Diseases caused by viruses, such as yellows, rosette, and phony, are especially prevalent on peaches, plums, cherries, apricots, and almonds. *See* PLANT VIRUSES AND VIROIDS.

The growth of fruit trees and the production of fruit may also be affected by unfavorable environmental conditions such as infertile soil, low temperatures, drought, and poor drainage. Eric G. Sharvelle

Citrus diseases. Citrus is subject to a great number of diseases, most of which are distributed throughout the citrus-producing countries of the world. Infectious diseases of citrus are caused by bacteria, fungi, and viruses and may be grouped according to the parts of trees affected, that is fruit, leaves and twigs, and trunks and roots.

Numerous rots, internal derangements, and external markings result from fungal and bacterial infections. Fruit decay develops chiefly during storage and transit from infections established largely in surface wounds during growing and from picking, packing, or marketing operations. Citrus fruit requires careful handling to avoid economic loss from decay caused by such fungi as the common green mold, *Penicillium digitatum;* blue mold, *P. italicum* (**Fig. 4**); brown rot, *Phytophthora citrophthora;* and cottony-rot, *Sclerotinia sclerotiorum*. Some decays are described as contact rots because the causal fungi can spread through an entire box of fruit from one infected fruit.

Some of the fruit-rotting fungi and others cause leaf-blight and dieback of twigs. A bacterium, *Xanthomonas citri*, is the cause of citrus canker which attacks leaves, twigs, and fruits. Introduced on citrus plants sent from Japan to Florida in 1913, it threatened to destroy the citrus industry there. Its eradication, requiring destruction by burning of 4,000,000

Fig. 4. Blue-green molds (*Penicillium italicum* and *P. digitatum*) which take a heavy toll in storage, transit, and market. The illustration shows how these decay-causing molds can spread by contact from a single infected citrus fruit. (*Photograph by L. J. Klotz*)

Fig. 5. Gummosis, a disease caused by the fungus *Phytophthra citrophthora*, which damages or kills many citrus trees. Infection by this fungus, a water mold, is favored by excessive moisture. (*Photograph by L. J. Kotz*)

orchard and nursery trees, is the most remarkable achievement in plant disease control.

Some of the fruit-rotting fungi, particularly species of *Phytophthora* and *Diplodia*, are responsible for root rots and gumming diseases of trees (**Fig. 5**).

Citrus is subject to several destructive virus diseases. One of these, psorosis, causes development of bark lesions and eventual death of trees. Since it was learned that psorosis is a bud-perpetuated disease that does not spread naturally, it has been possible to avoid it by propagating from psorosis-free sources. Another virus disease, tristeza, causes rapid decline or death of trees of sweet orange, grapefruit, and tangerine propagated on certain susceptible rootstock varieties. Tristeza virus is known to be spread by three species of aphids. It virtually destroyed the citrus industries of Argentina and Brazil in the 1930s. Damage from tristeza in the United States has been less because more of the citrus is on resistant rootstocks and the two species of the genus *Aphis* that spread it here are less efficient vectors than the one species common in South America. James M. Wallace

Fucales

A large and diverse order of conspicuous brown algae (Phaeophyceae) including such well-known seaweeds as *Fucus* and *Sargassum*. Definitive features include apical growth, production of sex organs in cavities, and a life history with only one somatic phase, which is diploid.

General structure. The mature thallus consists typically of a holdfast from which arise several axes with dichotomous or radial primary branching (**Fig. 1***a*). Growth is initiated by an apical meristem that results from the stimulation of neighboring cells by one or more apical cells located in a depression at the tip of each branch. An apical cell has lateral and basal cutting faces. Lateral derivatives of an apical cell continue to divide, producing a thin, pigmented, surface layer (meristoderm) and a thicker, colorless layer (cortex). The innermost lateral derivatives and the basal derivatives give rise to the medulla which, because of stretching, becomes filamentous at maturity. Vegetative thalli range in size, shape, and complexity from a small top-shaped structure, as in *Himanthalia*, to densely branched axes several meters long, differentiated into stems, leaves, and flotation organs (bladders or vesicles), as in *Sargassum*.

Reproduction. Sexual reproduction is oogamous, with oogonia and antheridia produced in flask-shaped cavities called conceptacles (Fig. 1*b*, *c*). Usually the conceptacles are localized at the swollen tips of vegetative branches or in special branches that range from long, dichotomously branched straps emerging from the vegetative thallus (*Himanthalia*) to small clusters of slender cylindrical axes borne in the axils of leaves (*Sargassum*).

The thallus in Fucales is diploid, with meiosis taking place within the sex organs. From one to eight eggs are produced in each oogonium, and usually 64 antherozoids (sperm) in each antheridium. Oogonia and antheridia occur on the same thallus or on different thalli. The antherozoid, like all motile reproductive cells in Phaeophyceae, is biflagellate but in some genera it departs from the usual situation by having the smooth, posteriorly directed flagellum longer than the hairy, anteriorly directed flagellum. Numerous antherozoids attempt to penetrate an egg, but the initial entry triggers the immediate formation of a fertilization membrane, which inhibits further entry. The union of gametes is followed by cell divisions that occur in a fairly precise order and result in a free-living embryo. The development of the embryo, which is easily observed and controlled in culture, has been studied extensively to provide information on the establishment of polarity by environmental factors. At the tip of the embryo there is a tuft of hairs, each with a meristematic basal region. The upper portion of one or more of these hairs abscises, while the lowermost meristematic cell metamorphoses into an apical cell.

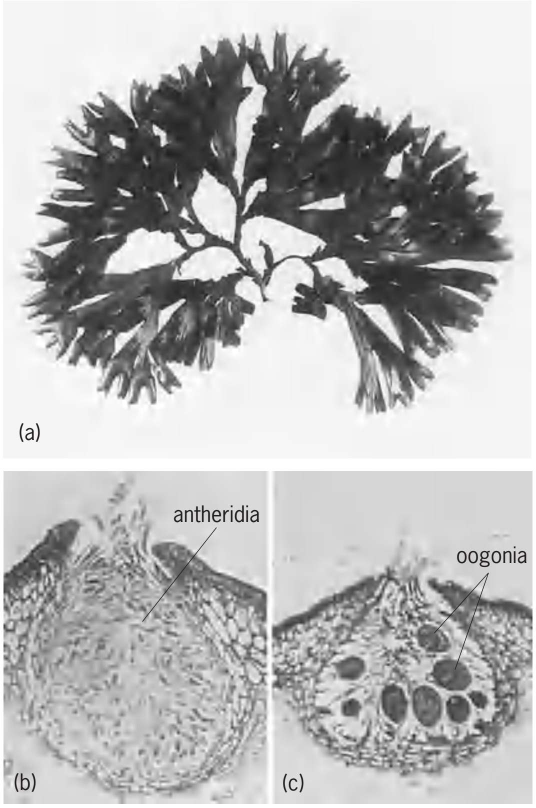


Fig. 1. ***Fucus*, a typical member of the order Fucales. (*a*) Plant with basal disk or holdfast and dichotomously branched frond with inflated tips containing reproductive cavities or conceptacles. (*b*) Section through male conceptacle lined with branched hairs bearing antheridia. (*c*) Section through female conceptacle lined with hairs and oogonia. (*Parts b and c from H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)**

The fucalean life history has been interpreted in two ways. In the first interpretation, the events are taken at face value: the thallus is considered a diploid gametophyte. In the second interpretation, the life history is compared to that of higher plants, in which the gametophytes are reduced to the cytological level. The oogonia and antheridia are considered megasporangia and microsporangia, the nuclei resulting from meiosis are considered megaspores and microspores, while subsequent nuclear divisions leading to the gametes supposedly represent gametophytes that have been lost in the course of evolution. Support for this interpretation is given by the observation of walls within the oogonium in certain genera.

Important families and species. In the family Fucaceae (rockweeds), the thallus typically is composed of a discoid or conical holdfast with a cluster of erect fronds that are branched dichotomously in one plane. In most genera, conceptacles are borne on swollen tips of branches. *Fucu* is common throughout the North Atlantic and North Pacific. Vesicles, when present, are often paired, with one on either side of the midrib. *Pelvetia*, similar to *Fucus* but with segments that are narrower, without a midrib, and usually without vesicles, is also common in the Northern Hemisphere, but is not found on the American side of the Atlantic. *Ascophyllum* occurs abundantly in the North Atlantic, both in exposed sites and in estuaries far from the open sea. It is characterized by large vesicles on the midline of the axis, and special swollen reproductive branches borne laterally. Loose, pinnate branching is superimposed on the primary dichotomous branching.

In subtropical and tropical regions of all oceans, the family Sargassaceae is common, dominating the biomass of benthic seaweeds (**Fig. 2**). The holdfast is discoid, conical, hapteroid (with rootlike outgrowths), or rhizomatous. The primary axis is condensed or percurrent (running from base to apex) and bears radially arranged secondary axes, which in turn produce radially or bilaterally arranged laterals. The primary axis is usually perennial, while the secondary axes are renewed annually. Vegetative branches as well as fertile branchlets are borne in the axils of subtending leaves.

Fig. 2. ***Sargassum*****, a member of the family Sargassaceae. Plant shows percurrent axis, dentate leaves, spherical vesicles or bladders, and fertile branchlets borne in axils of leaves.**

Sargassum is by far the largest and most important genus in the Sargassaceae. Hundreds of species and varieties have been described. Its taxonomy is complicated by morphological variation related to season, age, and habitat. In that portion of the North Atlantic between the West Indies and the Azores, an abundance of seaweed floats on relatively calm water. This region is called the Sargasso Sea because *Sargassum* is the primary constituent of the floating vegetation. The species of *Sargassum* in the Sargasso Sea reproduce solely by fragmentation.

Perhaps the best-known species of *Sargassum* is *S. muticum*, which is modest and well behaved in its native land (Japan) but noxiously intrusive elsewhere. It is a weed in the sociological sense of a plant growing outside its original range and in such a manner as to be unwanted. It was introduced to northwestern North America, presumably in shipments of oyster spat, just before World War II, and now occurs in protected waters from British Columbia to Baja California. It was detected in southern England in 1973 and has subsequently spread to nearby shores on the European continent. There is fear that it will eventually displace native species, and the long dense fronds are known to interfere in the operation of small boats in harbors.

In *Turbinaria*, a much smaller genus of Sargassaceae, the thallus is constructed on the same basic plan as *Sargassum*, but it is condensed into a structure resembling a gymnosperm cone. The leaves are modified into top-shaped or obpyramidal structures that partly conceal the fertile branchlets, like cone scales concealing seeds.

The family Cystoseiraceae, closely related to the Sargassaceae, is found in all temperate waters, but is especially well represented in southern Australia, where a consortium of members forms a nearly continuous belt in the sublittoral fringe.

The family Hormosiraceae comprises the single genus *Hormosira*, with probably only one species (*Hormosira banksii*) represented by several growth forms. This alga is common in shallow waters of New Zealand and the cool-temperate portion of Australia. The thallus has a striking appearance, being composed of swollen segments arranged like beads in a necklace. The conceptacles are scattered over the surface of the segments.

Rockweeds are well known to coastal peoples of the temperate regions of the Northern Hemisphere. Traditionally, the seaweed was either plowed into the soil as a green manure or burned to an ash that was spread on the fields. In modern chemical industry, *Ascophyllum* is an important source of alginate, especially in Norway. In a pulverized form, it is added to livestock and poultry feed for nutritive purposes. On the east coast of the United States, it is used to pack bait worms for shipment, and clumps are frequently encountered in San Francisco Bay, where they have been discarded from fishing boats, but so far the alga has not established itself there. Rockweeds are important ecologically because they protect a host of intertidal animals from desiccation during low tides. In warm waters, *Sargassum* is an especially significant constituent of the ecosystem because of its abundance and biomass. It is a minor source of alginate. *See* ALGAE; ALGINATE; PHAEOPHYCEAE. Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1997; J. B. Jensen, Morphological studies in Cystoseiraceae and Sargassaceae (Phaeophyceae) with special reference to apical organization, *Univ. Calif. Publ. Bot.*, 68:1–61, 1974.

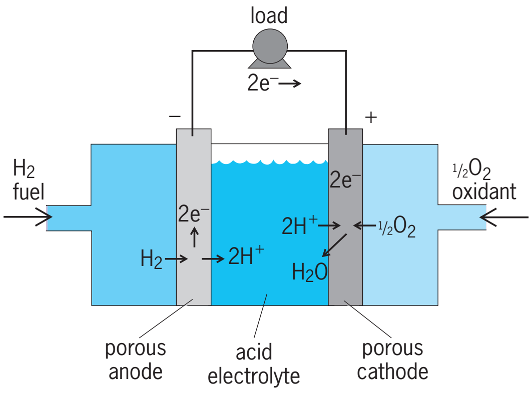
Fuel cell

An electrical cell that converts the intrinsic chemical free energy of a fuel directly into direct-current electrical energy in a continuous catalytic process. As in the classical definition of catalysis, the fuel cell itself should not undergo change; that is, unlike the electrodes of a battery, its electrodes ideally remain unchanged. For most fuel–oxidant combinations, the available free energy of combustion is somewhat less than the heat of combustion. In a typical thermal power conversion process, the heat of combustion of the fuel is turned into electrical work via a Carnot heat-engine cycle coupled with a rotating electrical generator. Since the Carnot conversion rarely proceeds at an efficiency exceeding 40% because of heat source and sink temperature limitations, the efficiency of conversion in a fuel cell can be greater than in a heat engine, especially in small devices. *See* CARNOT CYCLE; CATALYSIS.

Principle of operation. The fuel cell reaction usually involves the combination of hydrogen (H) with oxygen (O) [reaction (1)], as shown in the **illustration**.

$$H_2(g) + {}^{1}/{}_{2}O_2(g) \rightarrow H_2O(l) \qquad (1)$$

Under standard conditions of temperature and pressure, 25°C (77°F) and 1 atm (100 kilopascals), the reaction takes place with a free-energy change $\Delta G = -237$ kilojoules per mole of liquid water. Since the formation of water involves two electrons, this value corresponds to -1.23 electronvolts (1 eV = 96.5 kJ/equivalent). Thus, at thermodynamic equilibrium (zero current), the cell voltage should theoretically be 1.23 V, yielding a theoretical efficiency based on the higher heating value heat of combustion [ΔH for $H_2O(l) = -1.48$ eV] of 83.1%. If gaseous water at 100°C (212°F) is the product, the corresponding ΔG and lower-heating value ΔH are -1.16 eV and -1.25 eV. *See* FREE ENERGY.



Principle of operation of a fuel cell. (*After A. J. Appleby and F. R. Foulkes, Fuel Cell Handbook, Van Nostrand Reinhold, 1989*)

At a net (nonzero) current, all cells show losses in cell voltage (V). In low-temperature fuel cells, these are due largely to the kinetic slowness (irreversibility) of the oxygen reduction reaction, which requires the breaking of a double bond with transfer of four electrons per molecule in a complex sequence of reactions. Such cells show deviations from the theoretical 1.23-V open circuit. Typical values are 0.9 to 1.0 V. In high-temperature systems, oxygen reduction losses are less significant, since the reaction rate increases with temperature. However, the available free energy then decreases, falling to a value corresponding to about 1.0 V at 1000°C (1832°F). A further thermodynamic loss results from high cell fuel (or oxidant) conversion to avoid waste, so that the effective reversible potential is displaced from the standard state. At high temperature, the major loss is thermodynamic, which tends to compensate for the irreversible oxygen electrode losses at low temperature. As a result, cell voltages under typical loads vary from about 0.65 V for simple terrestrial cells to 0.9 V for aerospace cells. Cell voltage falls with increasing current per unit area. Since cell thermal efficiency based on the higher heating value of hydrogen (HHV, liquid water product) is given by $V/1.25$, cell performance is a compromise between relative cost (that is, kilowatts available per unit area) and fuel efficiency, to give the lowest cost of electricity for a given application. *See* OXIDATION-REDUCTION.

Fuels. While any chemically suitable fuel, including metals such as lithium (Li), sodium (Na), aluminum (Al), and zinc (Zn), may be used in a fuel cell, hydrocarbons (for example, natural gas) will not react at a significant rate in low-temperature fuel cells. They will crack thermally before reacting electrochemically if injected directly into high-temperature fuel cells. Practical hydrogen-bearing fuels have reactivities in low-temperature cells that increase in the order hydrocarbons $\ll$ ammonia $<$ methanol $\ll$ hydrazine $<$ hydrogen, provided that suitable anode catalysts are used. Simple low-power units operating directly on methanol at ambient temperature do find some use, and liquid-fueled hydrazine cells have also found specialized applications. However, the high manufacturing energy requirement for hydrazine, together with its high cost and hazardous nature, leaves hydrogen the only suitable general high-performance fuel candidate. *See* AMMONIA; HYDRAZINE; HYDROCARBON; HYDROGEN; METHANOL.

As oxidant, oxygen from air or as a cryogenic liquid for use in space is the first choice, although liquid oxidants (for example, hydrogen peroxide or nitric acid) have occasionally been used in specialized applications.

Because of the low solubility of H_2 and O_2 in electrolytes, a large area of contact at the electrode–electrolyte interface is required to achieve high reaction rates (0.2–0.8 A/cm^2 at atmospheric pressure, depending on the technology, and up to several amperes per square centimeter at elevated pressure). The large contact area is required because the reactions take place at the electrode–electrolyte interface. This large area is obtained by using porous catalytic materials with high surface areas that provide contact between electrolyte and gas, catalyze the reaction, maintain the electrolyte as a very thin layer at the surface of the electrode, and conduct electrons at the external circuit.

For practical fuel cells, hydrogen can be produced from readily available fuels, such as clean light distillate (for example, naphtha) usually by steam reforming, or from coal via gasification at high temperature (the direct use of coal or carbon has been abandoned). In the high-temperature cells under certain conditions, internal steam reforming of simple hydrocarbons and alcohols (for example, methane and methanol) can take place by the injection of the fuel with steam, which avoids cracking. Since methanol fuel reacts only slowly at low temperature, it is also steam-reformed to hydrogen. Methanol reforming takes place at only about 350°C (660°F), giving mixtures of hydrogen and carbon dioxide (CO_2) with a small amount of carbon monoxide (CO). In contrast, steam reforming of higher-molecular-weight alcohols or clean light distillates requires temperatures in excess of 700°C (1290°F). This favors mixtures of hydrogen and carbon monoxide, as in coal synthesis gas. *See* COAL GASIFICATION; SYNTHETIC FUEL.

Classification. Since the electrolyte that is used determines the chemical environment, hence stability, of the electrodes and other components, fuel cells are conveniently classified by electrolyte. There are three broad groups: aqueous, molten salt, and high-temperature solid ionic conductor.

Aqueous. To avoid excessive polarization, the major conducting ion in the electrolyte must be produced in the cell reaction at one electrode and must be consumed in the other. A good example is the aqueous alkaline fuel cell, using potassium hydroxide (KOH) as the electrolyte, with electrode reactions (2) and (3). In these reactions, the hydroxide ion (OH^-) is

$$2H_2 + 4OH^- \rightarrow 4H_2O + 4e^- \quad \text{(anode)} \qquad (2)$$

$$O_2 + 2H_2O + 4e^- \rightarrow 4OH^- \quad \text{(cathode)} \qquad (3)$$

the conducting ion. In acid cells, the conducting ion is hydrogen ion (H^+), produced from H_2 at the anode and consumed to give water at the cathode. Alkaline cells are capable of very high performance, and they have a benign corrosion environment that can allow the use of inexpensive (non-noble-metal) cathode catalysts. However, they cannot tolerate carbon dioxide and are not suitable for direct use with reformed carbonaceous fuels. Because of their efficiency, they have been used in space, operating at 260°C (500°F) with sintered nickel oxide cathodes used in the Apollo Program, and at about 85°C (185°F) in the high-performance lightweight system in the space shuttle orbiter. Both used cryogenic hydrogen and oxygen supplied at about 4 atm (400 kPa) pressure. The orbiter cell uses gold alloy cathodes and platinum alloy anodes bonded with nonwetting Teflon to maintain a stable interface. Low-capital-cost alkaline systems operating on "green" hydrogen were attracting attention in 2005.

Acid cells reject carbon dioxide and hence are suitable for use on hydrogen–carbon dioxide mixtures. However, carbon monoxide is an anode catalyst poison above certain critical levels that depend on cell operating temperature and anode catalyst composition. Unlike alkaline cells, acid cells require noble-metal catalysts, though they can be supported economically on nano-size carbon black or graphitic particles in low loadings (3 g or 0.1 oz per kilowatt or less). A high-performance system (the polymer electrolyte membrane or proton exchange membrane fuel cell, PEMFC) that has found favor for small units uses a stable fluorinated sulfonic acid polymer as the electrolyte. It is normally used with pure hydrogen fuel. Units vary in size from a few watts (operating directly on methanol as fuel) for portable power applications to greater than 100 kW (for experimental vehicle propulsion using pure hydrogen fuel). They operate at temperatures up to about 70–80°C (160–175°F), limited by the requirement for the presence of liquid water to ensure efficient proton conduction in the electrolyte. This means that they have very low (parts per million) carbon monoxide tolerance.

Throughout the 1990s, the most important acid fuel cell was the phosphoric acid fuel cell (PAFC) that used phosphoric acid as the electrolyte at 190–205°C (375–400°F), with Teflon-bonded low-loading platinum-based electrodes and graphite intercell separator plates. The high operating temperature was possible because self- ionizing phosphoric acid does not require liquid water for proton conduction. It made the anode catalysts tolerant to up to 1.5% CO input, simplifying the cleanup of the hydrogen-carbon dioxide feedstocks produced by steam reforming. It also simplified the cooling system, giving waste heat for cogeneration, and produced steam for reforming so no extra fuel was required for this purpose. This integrated system approach increased overall efficiency. Exhaust emissions levels are very low, allowing the power plant to be used where required, even in cities. *See* COGENERATION; ELECTRIC POWER GENERATION.

Molten salt. The only fuel-cell technology using molten salt electrolyte is the molten-carbonate fuel cell (MCFC), using molten carbonate (typically Li-K or Li-Na) in the temperature range 600–700°C (110–1300°F). The cells require that gaseous CO_2, recycled from the anode, exit to the cathode, as shown

Cell and system thermal efficiencies for different fuel-cell technologies and operating conditions

Type	Temperature, °C (°F)	Pressure, atm (kPa)	Current density, mA/cm^2	Voltage, V	System thermal efficiency (ac output/natural gas HHV), %
Alkaline	70 (158)	1 (100)	200	0.8	32
Phosphoric acid	195 (374)	1 (100)	330	0.65	36
PEM	75 (167)	1 (100)	400	0.7	35
Molten carbonate (internal reforming)	650 (1202)	1 (100)	160	0.78	45.6
Solid oxide (integrated reforming)	1000 (1832)	1 (100)	200	0.68	43.3

in the electrode reactions (4) and (5). MCFCs are attractive because they use relatively inexpensive materials. They are particularly suited for use with reformed hydrocarbon fuels, since they can consume CO directly via the internal water–gas shift process;

$$H_2 + CO_3{}^{2-} \rightarrow H_2O + CO_2 + 2e^- \qquad (4)$$

$$CO_2 + {}^1/_2O_2 + 2e^- \rightarrow CO_3{}^{2-} \qquad (5)$$

and they can be designed to reform natural gas internally if the stack (fuel cells connected in series) gas-flow channels incorporate a reforming catalyst. This has the engineering advantage of adsorbing 40% of the stack waste heat, greatly simplifying cooling by reactant gas streams. This is the principle of the direct fuel cell (DFC). Cell and system performance is higher than that of the PAFC, and they may attain over 50% system efficiency, or higher values, if they use a gas or steam turbine bottoming cycle powered by their high-quality waste heat. *See* COGENERATION.

MCFCs were first developed as stacks in the early 1980s and were scaled up to the 100-kW demonstrator stage by 1989. During the 1990s, a 2-MW DFC unit was demonstrated. Two other 2-MW units and a 1-MW unit operating on digester gas followed. The standard product is a 250–300-kW DFC unit with 0.78 m^2 (8.4 ft^2) active area cells.

High-temperature solid ionic. The solid-oxide fuel cell (SOFC) operates between 800 and 1000°C (1470 and 1830°F) and uses thin-layer yttria-stabilized zirconia as the electrolyte. Conducting oxides matched for thermal expansion are used as the cathode and to interconnect the cells. The anode is a nickel cermet. One such system uses tubular 2 m long by 2.2 cm diameter cells, 1.5 m in active length (834 cm^2 active area). The cells are made from porous cathode material and are supplied by air through a central alumina tube that acts as a ceramic heat exchanger. Partially reformed natural gas and steam are passed between the cells from the direction of the closed ends. Cells are linear-interconnected to the inner cathode extending through the electrolyte and anode, and insulated from the latter. The surface of the interconnect is connected to the anode of the next tube via a nickel felt. *See* CERMET.

Unused reformate (15%) burns with unused air in an alumina plenum chamber at the top. The system offers high-quality waste heat. Internal reforming at the operating temperature is so rapid that too much cell cooling occurs, so separate integrated prereforming is employed, using part of the waste heat. The remainder can be used in cogeneration, or in a 3-atm (300-kPa) pressurized system the heat can be supplied to a turbine, which drives the compressor and produces additional power. Scale-up to 250 kW took place in the 1990s. Several 250-kW units, including one pressurized system, have been delivered.

Elsewhere, researchers are concentrating on planar cells, in which one major problem is sealing. Research is aimed at reducing the SOFC operating temperature by the use of thinner components and/or the use of different families of conducting materials. This approach will allow the use of metal heat exchangers and treated metal bipolar plates.

Performance. The **table** shows cell and system thermal efficiencies for different fuel-cell technologies and operating conditions using hydrogen derived from natural gas via steam reforming. System efficiencies are based on the higher heating value (HHV) of natural gas and include parasitic power losses and dc/ac conversion. The HHV is used because the gas is purchased based on this value. For rotating machinery, the lower heating value (LHV) efficiencies are usually used. These are a factor of 1.1 higher. The thermal efficiency of the alkaline system is low because pure hydrogen must be separated and fuel must be burned for reforming. The two high-temperature systems are assumed to have no bottoming cycle. Assumptions used are SOFC fuel utilization 85%, MCFC 78%, inverter efficiency 93%, with auxiliaries 7% of ac output. They reflect the efficiency of heat recovery in reforming and upgrading methane ($\Delta H = -1.04$ eV for gaseous water product) to hydrogen ($\Delta H = -1.25$ eV) internally within the cell and system itself. These show the concern with fuel-cell power-system design, rather than simply emphasizing the fuel cell itself. In practice, systems would probably be operated at higher current density and lower efficiency to reduce capital cost and, therefore, the cost of electricity. A. John Appleby

Bibliography. A. J. Appleby, Fuel cell technology, status and future prospects, *Energy*, 21(7–8):521–653, Elsevier Science, Amsterdam, 1996; A. J. Appleby and F. R. Foulkes, *Fuel Cell Handbook*, Van Nostrand Reinhold, 1989; R. L. Busby, *Hydrogen and Fuel Cells: A Comprehensive Guide*, PennWell, 2005; P. Hoffman, *Tomorrow's Energy, Hydrogen,* *Fuel Cells, and the Prospects for a Cleaner Planet*, MIT Press, 2002; K. Kordesch and G. Simader, *Fuel Cells and Their Applications*, Wiley-VCH, 1996; J. Larminie and A. Dicks, *Fuel Cell Systems Explained*, 2d ed., Wiley, Chichester, U.K., 2003.

Fuel gas

A fuel in the gaseous state whose potential heat energy can be readily transmitted and distributed through pipes from the point of origin directly to the place of consumption. The development and use of fuel gases is closely associated with the progress of civilization. Such gases have become prominent in the industrial development period since 1900.

The types of fuel gases are natural gas, LP gas, refinery gas, coke oven gas, and blast-furnace gas. The last two are used in steel mill complexes. Typical analyses of several fuel gases are presented in the **table**. Since these analyses are based on dry gases, the heating value of gases saturated with water vapor would be slightly lower than the values shown.

Most fuel gases are composed in whole or in part of the combustibles hydrogen, carbon monoxide, methane, ethane, propane, butane, and oil vapors and, in some instances, of mixtures containing the inerts nitrogen, carbon dioxide, and water vapor.

Natural gas. The generic term “natural gas” applies to gases commonly associated with petroliferous geologic formations. As ordinarily found, these gases are combustible, but nonflammable components such as carbon dioxide, nitrogen, and helium are often present. Natural gas is generally high in methane. Some of the higher paraffins may be found in small quantities.

The olefin hydrocarbons, carbon monoxide, and hydrogen are not present in American natural gases. The term “dry natural gas” indicates less than 0.1 gal (1 gal, U.S. = $0.003785\ \text{m}^3$) of gasoline vapor occurs per 1000 ft^3 (1 $\text{ft}^3 = 0.028\ \text{m}^3$); “wet natural gas” indicates more than 0.1 gal/1000 ft^3. “Sweet” and “sour” are terms that indicate the absence or presence of hydrogen sulfide.

There is no single composition which might be termed typical natural gas. Methane and ethane constitute the bulk of the combustible components; and CO_2 and nitrogen, the inerts. The net heating value of natural gas served by a utility company is often 1000 to 1100 Btu/ft^3 (3.8 to 4.2×10^7 joules/m^3).

Natural gas is an ideal fuel for heating because of its cleanliness, ease of transportation, high heat content, and the high flame temperature. *See* LIQUEFIED NATURAL GAS (LNG); NATURAL GAS.

LP gas. This term is applied to certain specific hydrocarbons, such as propane, butane, and pentane, which are gaseous under normal atmospheric conditions but can be liquefied under moderate pressure at normal temperatures. *See* LIQUEFIED PETROLEUM GAS (LPG).

Oil gas. This term encompasses a group of gases derived from oils by exposure of such oils to elevated temperatures. Refinery oil gases are those obtained as by-products during the thermal processing of the oil in the refinery. They are used primarily for heating equipment in the refinery. Gas made by thermal cracking of oil was formerly very important as an urban fuel gas, but has been almost totally displaced by natural gas, and the equipment for its production dismantled. A typical oil gas consists of saturated and unsaturated hydrocarbons and has a heating value of 1300 to 2000 Btu/ft^3. Methane, ethane, propane, butane, ethylene, and propylene are the main constituents.

Coal gas. Until about 1940, gas produced from coal was an important part of the energy mix in the United States. These gases were rapidly replaced by natural gas in the distribution systems of gas utilities serving the residential, commercial, and industrial markets. During 1971 it became clear that the supply of natural gas would probably be insufficient to meet the ever-increasing demand. Therefore, it is expected that gas will once again be produced from coal. The older processes will be replaced by modern methods capable of very high production rates at improved efficiency and reduced cost.

Coal gasification can be accomplished in a large number of ways, including pyrolysis or partial oxidation with air or oxygen and steam. Various processes operate as fixed beds or fluidized beds or with the coal entrained. The pressure may vary from near atmospheric to 1000 pounds per square inch gage (1 lb/in.^{2} = 6895 pascals) or more. All operate at high temperatures (1200–3000°F; 650–1645°C). The direct products of gasification vary in heating value from 120 to 150 Btu/ft^3 or 4.6 to 5.7×10^3 J/m^3

Typical gas analyses*

	Analysis, % vol									Btu/ft^3†	
Type	CO_2	O_2	CO	H_2	CH_4	C_2H_6	C_3H_8 and C_4H_{10}	N_2	Specific gravity	Gross	Net
Dry natural gas	0.2				99.2			0.6	0.56	1007	906
Propane (LP)						2.6	97.3	0.1	1.55	2558	2358
Refinery oil gas		0.2	1.2	6.1	4.4	72.5	15.0	0.6	1.00	1650	1524
Coke oven gas	2.0	0.3	5.5	51.9	32.3		3.2	4.8	0.40	569	509
Blast-furnace gas	11.5		27.5	1.0				60.0	1.02	92	92
Producer gas	8.0	0.1	23.2	17.7	1.0			50.0	0.86	143	133

*From L. Shnidman (ed.), *Gaseous Fuels,* 2d ed., American Gas Association, 1954.
†1 Btu/ft^3 = 3.8×10^4 joules/m^3.

(low-Btu gas), to 300 Btu/ft^3 or 11.4×10^7 J/m^3 (medium-Btu gas or synthesis gas), to as high as 600 Btu/ft^3 or 2.28×10^7 J/m^3. Low-Btu gas can be used as fuel for industrial processes or for production of electrical power by electric utilities. Medium-Btu gas, which consists principally of carbon monoxide and hydrogen, can be used directly as a fuel, or it can be upgraded by catalytic methanation to essentially pure methane, which is, for all practical purposes, identical to natural gas. Such a product gas is commonly called pipeline-quality gas or substitute natural gas (SNG). *See* COAL GASIFICATION.

Low-Btu gas. Producer gas is made by oxygen-deficient combustion of coal or coke, in which process a mixture of air and steam is blown upward through a thick hot bed of coal or coke. The gas is high in nitrogen introduced in the air. Its heating value is low, its specific gravity is high, and the percentage of inerts is high. Producer gas, which contains 23–27% carbon monoxide, is used as it comes from the generators after some preliminary purification. It was once the cheapest form of industrial gas, and could again become important; it may be possible to produce it in modernized equipment on a large scale.

Blast-furnace gas is a by-product from the manufacture of pig iron. Like producer gas, it is derived from the partial combustion of coke. Some of the combustibles in the gas are used to reduce the iron ore; thus the final gas contains about 27% carbon monoxide and more than 70% of inert gases (CO_2 and N_2), giving it the lowest heating value, less than 100 Btu/ft^3 or 3.8×10^6 J/m^3, of any usable fuel gas. It is used for the operation of gas engines, heating by-product coke ovens, steel plant heating, steam raising, and crude heating. *See* FOSSIL FUEL; PYROLYSIS.

Medium-Btu gas. Coke oven gas is the only important gas that has an intermediate heating value. The gas is made by destructive distillation of a packed bed of coal out of contact with air. The process results in the formation of coke, which is used in the blast furnace. The gas is utilized totally within the steel-making complex.

The combination of the 1974 oil embargo, the energy crisis, and restrictions on air pollution resulted in the proposal of a large number of modern processes. Nearly all these processes use a mixture of steam and oxygen to combust coal. Less oxygen than is required for complete combustion is used; therefore, the products are primarily carbon monoxide and hydrogen. In some cases, depending on the end use of the gas, moderate amounts of methane and liquid products are produced. *See* COKE.

SNG. Gas which is to be used within the modern natural gas transmission and distribution system must be essentially methane. The medium-Btu gas processes can be used as precursors to produce substitute natural gas (SNG), as noted above. Other new processes which operate at very high pressure (1000 lb/in.^{2} gage) are being developed to produce very large quantities of gas (2.5×10^8 ft^3/day of 1000 Btu/ft^3 gas or 7×10^6 m^3/day of 3.8×10^7 J/m^3 gas) to supplement the declining supplies of natural gas. Some of these processes use the partial combustion of oxygen as a direct source of heat, and others use air combustion in a variety of indirect modes.

Nonfossil sources. In view of the immediate shortage of natural gas and the longer-range (Af5D. 2030–2060) forecast that all fossil fuels will be consumed, there are several projects which may ensure a perpetual supply of fuel gas. These include conversion of waste materials to gas by pyrolysis or by anaerobic digestion. Large efforts have been initiated to convert solar energy to fuel gas. This would be accomplished by growing plants on land or sea and converting the harvest to gas in processes similar to those used on wastes. Finally, there has been a large effort to develop a "hydrogen economy." Hydrogen would be produced by thermochemical processes using heat from various sources such as atomic energy reactors.

Jack Huebler

Bibliography. M. Crow et al.,*Synthetic Fuel Technology Development in the United States: A Retrospective Assessment*, 1988; R. A. Myers, *Handbook of Synfuels Technology*, 1984; R. F. Probstein and R. E. Hicks, *Synthetic Fuels*, 1990.

Fuel injection

The pressurized delivery of a metered amount of fuel into the intake airflow or combustion chambers of an internal combustion engine. Metering of the fuel charge may be performed mechanically or electronically. In a diesel engine, the fuel is injected directly into the combustion chamber (direct injection) or into a smaller connected auxiliary chamber (indirect injection). In the spark-ignition engine, the fuel is injected into the air before it enters the combustion chamber by spraying the fuel into the airstream passing through the throttle body (throttle-body injection) or into the air flowing through the port to the intake valve (port injection). *See* COMBUSTION CHAMBER.

Diesel fuel injection. The diesel engine must be supplied with fuel from the injection nozzle at a pressure of 1500–5000 lb/in.^{2} (10–35 megapascals) for indirect-injection engines, and up to 15,000 lb/in.^{2} (100 MPa) or higher for direct-injection engines. The high pressure is necessary to deliver fuel against the highly compressed air in the engine cylinders at the end of the compression stroke, and to break up the fuel oil which has low volatility and is often viscous. Extremely accurate fuel metering is necessary, with the start of injection occurring within a precision of up to 1° of engine crankshaft angle. A timing device in the injection pump automatically advances the start of of fuel delivery as engine speed increases to optimize the start of combustion.

The intake air is not throttled in a diesel engine, with load and speed control accomplished solely by controlling the quantity of fuel injected. The mean effective pressure developed by combustion is controlled by the volumetric capacity of the injection pump. To prevent an unloaded diesel engine from increasing in speed until it destroys itself, a governor is required to limit maximum engine speed. *See* DIESEL CYCLE; DIESEL ENGINE; GOVERNOR; INTERNAL COMBUSTION ENGINE; MEAN EFFECTIVE PRESSURE.

In a typical diesel fuel-injection system, a low-pressure supply pump delivers fuel from the fuel tank to the injection pump, which pressurizes, times, and meters the fuel charge and then delivers it through tubing to an injection nozzle that sprays the pressurized fuel into the combustion chamber. Either all fuel lines between the injection pump and nozzles are made of equal length, or a high-pressure manifold or fuel rail is used. The high fuel-injection pressure is produced by an engine-driven reciprocating piston pump fitted to very close tolerances, either in the injection-pump assembly or in a unit injector.

There are two types of injection pump: the in-line pump and the distributor pump. The in-line pump has two or more pumping elements that pressurize the fuel and may also time and meter it. Each pumping element, which is the combination of a pump plunger and its barrel, serves only one engine cylinder. As the plunger moves from the bottom to the top of its stroke, it closes off a spill port and begins to compress the fuel in the high-pressure chamber. The increasing pressure opens the delivery valve in the pump, allowing a pressure wave to travel at the speed of sound toward the injection nozzle. When the nozzle-opening pressure is reached, the nozzle needle opens against spring force, and then fuel is injected until the pump plunger opens the spill port. This relieves the pressure and the delivery valve closes.

In the distributor pump (**Fig. 1**), one pumping element is used to provide metered delivery to all cylinders. The plunger stroke delivers the fuel to a rotating distributing device, which then directs the fuel to the appropriate engine cylinder. The number of plunger strokes during one revolution of the injection-pump driveshaft corresponds to the number of engine cylinders.

Instead of an injection pump, some engines have a unit fuel injector for each cylinder. It receives fuel under supply-pump pressure and is then actuated by an engine or electronic mechanism to meter and inject the fuel charge into the combustion chamber at high pressure and at the proper time.

Gasoline fuel injection. On automotive spark-ignition engines, the carburetor has largely been replaced by a gasoline fuel-injection system with either mechanical or electronic control of fuel metering. Many of the systems are of the speed-density type, in which the mass airflow rate is calculated based on cylinder displacement and the measured intake-manifold absolute pressure (engine load), engine speed, intake-manifold air temperature, and theoretical volumetric efficiency. When the feedback signal from an exhaust-gas oxygen sensor is included, these systems allow the engine air-fuel ratio to be maintained near the stoichiometric ratio (14.7:1) during normal operating conditions. This minimizes exhaust emissions. *See* CARBURETOR.

In the typical gasoline fuel-injection system (**Fig. 2**), an electric fuel pump provides a specified fuel flow at the required system pressure to one or

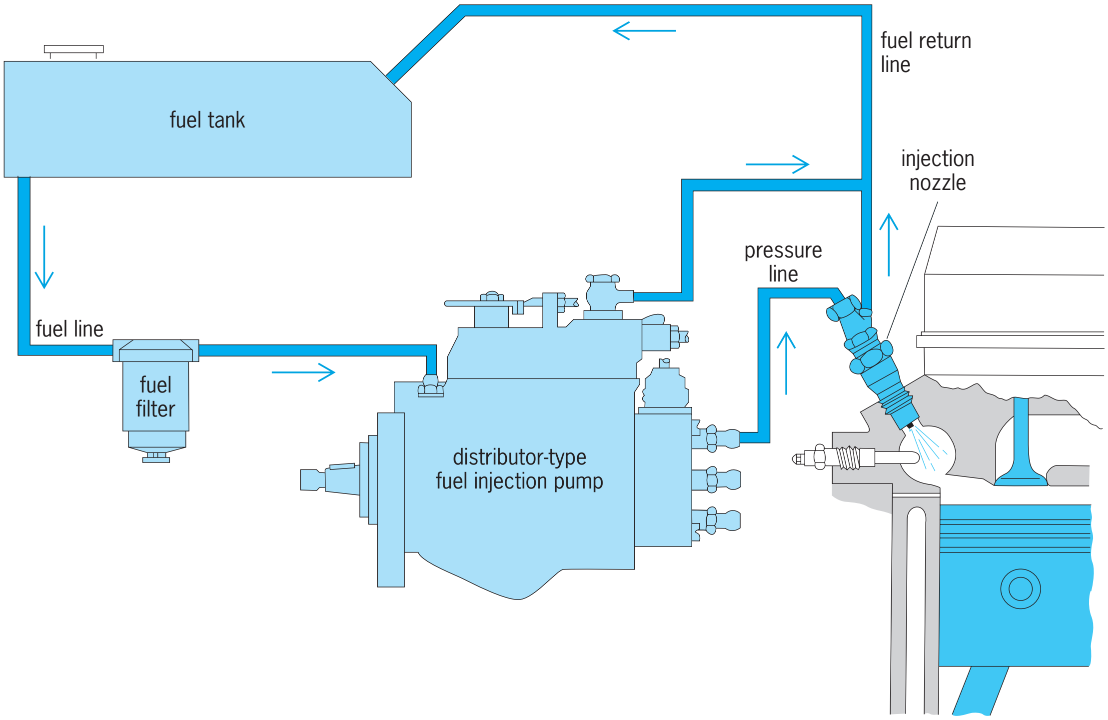


Fig. 1. Diesel-engine fuel-injection system using a distributor-type fuel-injection pump. (*Robert Bosch Corp.*)

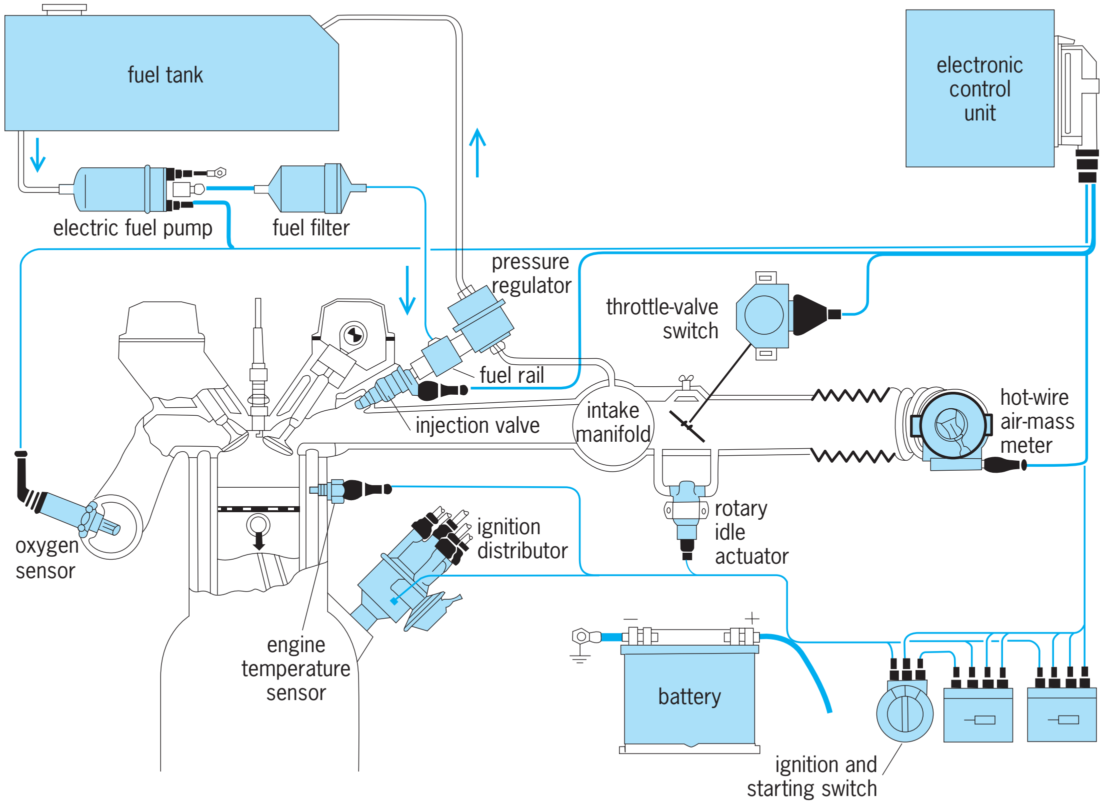


Fig. 2. **Electronic port fuel-injection system for a spark-ignition engine. (*Robert Bosch Corp.*)**

more fuel-injection valves, or fuel injectors. The gasoline fuel injector is an electromagnetic (solenoid-operated) or mechanical device used to direct delivery of or to meter pressurized fuel, or both. A fuel-pressure regulator maintains a controlled fuel pressure at each injector, or a controlled differential pressure across the injector.

The three types of gasoline fuel injection are throttle-body (central or single-point) injection, port (multipoint) injection (Fig. 2), and continuous injection. Throttle-body injection is an electronically controlled system in which one or two fuel injectors are located in the throttle body, usually above the throttle valves. The quantity of fuel delivered is determined by an electronic control unit whose output signal controls injector open time. This is based on information received by the electronic control unit from several sensors that monitor the operating conditions of the engine.

In electronic port injection systems, the fuel is delivered to a fuel manifold or rail that distributes it to an individual fuel injector located in the intake port for each cylinder. The injectors may be triggered either in groups or sequentially. A continuous injection system may be controlled either electronically or mechanically while fuel flows continuously from the injectors independent of cylinder events. Fuel flow rate is controlled by a variable orifice or fuel pressure. *See* FUEL PUMP; FUEL SYSTEM.

Donald L. Anglin

Bibliography. *Bosch Automotive Electric/Electronic Systems*, 1986; *Bosch Automotive Handbook*, 1986; W. H. Crouse and D. L. Anglin, *Automotive Fuel, Lubricating, and Cooling Systems*, 1981.

Fuel oil

Any of the petroleum products which are less volatile than gasoline and are burned in furnaces, boilers, or other types of heaters. The two primary classes of fuel oils are distillate and residual. Distillate fuel oils are composed entirely of material which has been vaporized in a refinery distillation tower. Consequently, they are clean, free of sediment, relatively low in viscosity, and free of inorganic ash. Residual fuel oils contain fractions which cannot be vaporized by heating. These fractions are black and viscous and include any inorganic ash components which are in the crude. In some cases, whole crude is used as a residual fuel.

Uses. Distillate fuel oils are used primarily in applications where ease of handling and cleanliness of combustion are more important than fuel price. The most important use is for home heating. Distillate fuel oils are also used in certain industrial applications where low sulfur or freedom from ash is important. Certain types of ceramics manufacture are examples. Increasing amounts of distillate fuel oils have been burned in gas turbines used for electricity generation. *See* ELECTRIC POWER GENERATION; GAS TURBINE.

Residual fuel oils are used where fuel cost is an important enough economic factor to justify additional investment to overcome the handling problems they pose. They are particularly attractive where large volumes of fuel are used, as in electric power generation, industrial steam generation, process heating, and steamship operation.

Fuel oils are frequently made for specific uses, and these may be distillates, residuals, or mixtures of the two. Therefore, terms such as domestic fuel oils, diesel fuel oils, and heavy fuel oils are being used in classification schemes.

Combustion. Fuel oils are burned efficiently by atomizing them into fine droplets, about 50–100 micrometers in diameter, and injecting the droplet spray into a combustion chamber with a stream of combustion air. Small home heating units use a pressure atomizing nozzle. Larger units use pressure, steam, or air atomizing nozzles, or else a spinning cup atomizer. Older units employ natural draft to induce combustion air, but new units usually use forced draft, which gives better control of air-to-fuel ratio and increased turbulence in the combustion chamber. This permits more efficient operation and clean, smoke-free combustion. *See* OIL FURNACE.

Tests and specifications. Fuel oils are blended to meet certain tests and specifications which ensure that they can be handled safely and easily, that they will burn properly, and that they will meet air-pollution regulations. These tests include (1) flash point, which determines that the fuel can be stored and handled without danger of explosion; (2) pour point, which determines that it will not solidify under normal handling conditions; (3) carbon residue, which determines that it will not coke during handling or form undue amounts of carbon particulates during combustion; (4) sediment, which determines that it will not clog pumps and nozzles; (5) sulfur content, which determines how much sulfur dioxide will be emitted during combustion; (6) ash content; and (7) viscosity or resistance to flow.

Viscosity. Viscosity is an important factor in determining the grade of a fuel oil. Distillate fuels and light grades of residuals have a low viscosity so that they can be handled and atomized without heating. Heavier grades have high viscosity and must be heated in order to be pumped and atomized. The grade classification of the fuel oil determines what viscosity range it is blended to, and therefore the temperature level required in the preheat system.

Ash content. Fuel ash content may be important in residual fuels, even though ash content is negligible compared with coal, because certain ash components may cause slagging and corrosion problems. Vanadium and sodium are frequently present in oil ash. During combustion, they are converted to materials which tend to accumulate on boiler and superheater tubes. Under certain conditions, they may cause corrosion of these tubes.

Sulfur content. Sulfur content of fuel oil is carefully controlled in order to meet air-pollution regulations. This has required a drastic reduction in sulfur content of residual fuel oils in many areas. To achieve the reduction, extensive processing and changes in blending procedure were required, which caused major increases in the price of fuel oil.

A detailed list of specifications for fuel oils is given in the *Annual Book of ASTM Standards*, published by the American Society for Testing and Materials, Philadelphia. *See* PETROLEUM PRODUCTS.

C. W. Siegmund

Fuel pump

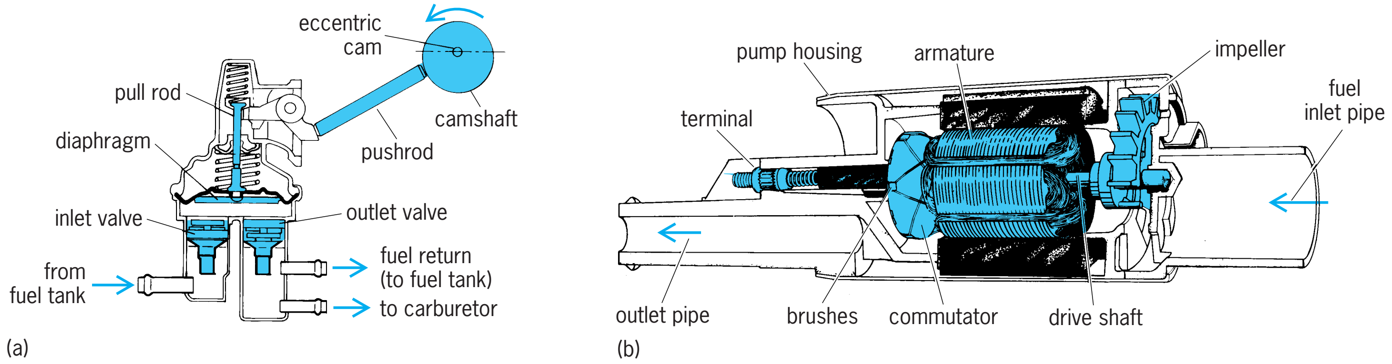
A mechanical or electrical pump for drawing fuel from a storage tank and forcing it to an engine or furnace. The type of pump chosen for a given fuel depends to a great extent on the volatility of the liquid to be pumped. In a gasoline engine the fuel is highly volatile at ambient temperature. Therefore, the fuel line is completely sealed from the tank to the carburetor or fuel-injection system to prevent escape of fuel and to enable the pump to purge the line of vapor in the event of vapor lock—a condition in which the fuel vaporizes owing to abnormally high ambient temperature. *See* CARBURETOR; FUEL INJECTION; FUEL SYSTEM.

Most carbureted gasoline engines use a spring-loaded diaphragm-type mechanical pump which is normally actuated by a rocker arm or pushrod that rides on an eccentric on the engine camshaft (**illus.** *a*). Electric motor–driven and solenoid-operated diaphragm pumps and plunger pumps are also available that can be mounted near the main fuel tank to minimize vapor lock in the fuel lines. *See* VAPOR LOCK.

Many gasoline-engine vehicles have a submersible electric fuel pump, which serves as the main supply pump, located in the fuel tank. In some fuel-injection systems, the in-tank pump is used as the supply pump for a high-pressure fuel-injection pump. The in-tank pump may be of the gear, plunger, sliding-vane, or impeller type (illus. *b*).

Diesel engines normally use a gear, plunger, or vane-type pump to supply fuel to the injection pump. In the diesel engine, where fuel is injected at high pressure through an injection nozzle into the highly compressed air in the combustion chamber, a plunger or piston serves as its own inlet valve and as the compression member of the injection pump. When the required high pressure is reached in the injection nozzle, a spring-biased needle valve opens and fuel sprays into the combustion chamber. In an oil-fired furnace, although nozzle pressures need not be so high as in diesel engines, a piston pump is also used to provide positive shutoff of the fuel line when the pump stops. *See* DIESEL ENGINE.

In a liquid-fueled rocket with pump feed, both fuel pump and oxidizer pump are typically centrifugal or turbine pumps, usually driven by an accessory turbine engine that burns a small quantity of pumped fuel and oxidizer. The pumps must withstand the corrosive liquids and have close clearances, preferably with no leakage. Flow rates may be limited by cavitation at the pump inlet, especially because the fuel or cryogenic fluid, such as liquid oxygen, may be



Fuel pumps for gasoline engines. (*a*) In-line mechanical diaphragm pump (*Chrysler Corp*). (*b*) In-tank electric impeller pump (*AC-Delco Division of General Motors Corp.*).

near its vapor pressure because of its high velocity. *See* CAVITATION.

Fuel pumps must operate before the engine or furnace that they serve can function. Consequently, an auxiliary or starter motor is usually an essential adjunct to a fuel pump. Fuel pumps should be easily accessible and mounted so that the outlet line can be disconnected and a pressure gage installed to check operating pressure. In diaphragm-type pumps, the diaphragm material must be capable of withstanding degradation by aromatic-type fuels. Donald L. Anglin

Bibliography. *Bosch Automotive Electrics/Electronics*, 1988; *Bosch Automotive Handbook*, 1986; W. H. Crouse and D. L. Anglin, *Automotive Fuel, Lubricating, and Cooling Systems*, 1981.

Fuel system

The system that stores fuel for present use and delivers it as needed to an engine; includes the fuel tank, fuel lines, pump, filter, vapor return lines, carburetor or injection components, and all fuel system vents and evaporative emission control systems or devices that provide fuel supply and fuel metering functions. Some early vehicles and other engines had a gravity-feed fuel system, in which fuel flowed to the engine from a tank located above it. Automotive and most other engines have a pressurized fuel system with a pump that draws or pushes fuel from the tank to the engine. *See* CARBURETOR; FUEL INJECTION; FUEL PUMP.

Automobile. The commonly used components for automobile and stationary gasoline engines are fuel tank, fuel gage, filter, electric or mechanical fuel pump, and carburetor or fuel-injection system (**Fig. 1**). In the past, fuel metering on automotive engines was usually performed by a carburetor. However, this device has been largely replaced by fuel injection into the intake manifold or ports, which increases fuel economy and efficiency while lowering exhaust gas emissions. Various types of fuel management systems are used on automotive engines, including electronically controlled feedback carburetors, mechanical continuous fuel injection, and sequential electronic fuel injection. *See* AUTOMOBILE.

Most carbureted engines use a diaphragm-type mechanical fuel pump that is operated by the rotation of an eccentric on the engine camshaft. Many gasoline fuel-injected engines use an impeller-type electric fuel pump located in the fuel tank (Fig. 1). This helps prevent problems with vapor lock, particularly during short engine stops on hot days, which results when the volatile fuel boils in the supply line and pumping chamber. *See* VAPOR LOCK.

Electric fuel pumps may have plunger-, diaphragm-, or impeller-type pumping elements, and are designed so that their speed of operation and current consumption depend upon the rate of fuel use. Safety devices prevent fuel delivery if the vehicle rolls over of if the engine stops running. *See* AUTOMOTIVE ENGINE. Donald L. Anglin

Aircraft. The presence of multiple engines and multiple tanks complicates the aircraft fuel system. Also, the reduction of pressure at altitude necessitates the regular use of boost pumps, submerged in the fuel tanks, which are usually of the centrifugal type and electrically driven. These supplement the engine-driven fuel supply pumps, which are usually of the gear or eccentric vane type.

Components of a typical aircraft fuel system include one main and two auxiliary tanks with their gages; booster, transfer, and engine-driven pumps; various selector valves; and a fuel jettisoning or defuel valve and connection, which is typical also of what would be needed for either single-point ground

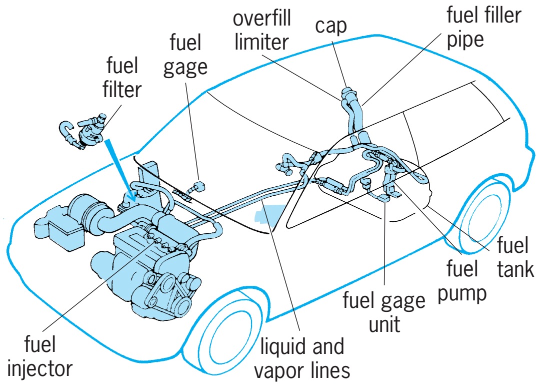


Fig. 1. Components of an automobile fuel system. (*Chrysler Corp.*)

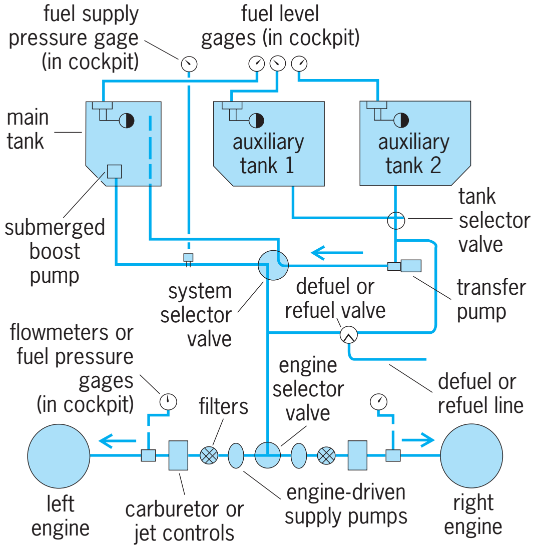


Fig. 2. Diagram of a typical aircraft fuel system.

or flight refueling (**Fig. 2**). The arrangement is usually such that all the fuel supply will pass to the engines by way of the main tank, which is refilled as necessary from the auxiliary tanks. In case of emergency, the system selector valve may connect the auxilary tanks to the engines directly. Tank vents, not shown, are arranged so that overflow will go safely overboard.

Because of the large quantities of fuel used, aircraft fuel systems are often contaminated with dirt, metal chips, and lint; adequate filtration is therefore essential to reliable service. *See* AIRCRAFT ENGINE.

J. A. Bolt

Fugacity

A function introduced by G. N. Lewis to facilitate the application of thermodynamics to real systems. Thus, when fugacities are substituted for partial pressures in the mass action equilibrium constant expression, which applies strictly only to the ideal case, a true equilibrium constant results for real systems as well.

The fugacity f_i of a constituent i of a thermodynamic system is defined by Eq. (1) [where μ_i is the

$$\mu_i = \mu_i^* + RT \ln f_i \qquad (1)$$

chemical potential and μ_i^* is a function of temperature only], in combination with the requirement that the fugacity approach the partial pressure as the total pressure of the gas phase approaches zero. At a given temperature, this is possible only for a particular value for μ_i^*, which may be shown to correspond to the chemical potential the constituent would have as the pure gas in the ideal gas state at 1 atm pressure. This definition makes the fugacity identical to the partial pressure in the ideal gas case. For real gases, the ratio of fugacity to partial pressure, called the fugacity coefficient, will be close to unity for moderate temperatures and pressures. At low temperatures and appropriate pressures, it may be as small as 0.2 or less, whereas at high pressures at any temperature it can become very large.

The fugacity concept is not restricted to gaseous systems, however. Because of its relation to the chemical potential, the basic thermodynamic criterion of equilibrium requires that the fugacity of a constituent have the same value at equilibrium for every phase in which it is present. This permits the indirect determination of the fugacity for a condensed phase through the calculation of the value for the equilibrium vapor phase, for which the fugacity may be computed routinely if the dependence of its volume on temperature, pressure, and composition is known. Thus, results are readily obtained for a pure gas, but because of the more extensive data required, accurate calculations have been made for very few mixtures.

For an ideal solution, the fugacity is given by the mole fraction of the constituent multiplied by the fugacity of the pure constituent at the temperature and pressure of the solution. For liquid solutions, this is the thermodynamic counterpart of Raoult's law, but the relation applies also to ideal gaseous solutions and can serve for the prediction of the properties of real gas mixtures. Where no equation of state data are available for pure gases, their fugacities can be estimated by means of the generalized fugacity coefficient chart. At sufficiently high dilution in liquid solution, the fugacity of a nondissociating solute will become proportional to its concentration, the proportionality constant depending on the concentration scale used; this is the thermodynamic statement of Henry's law.

The dependence of the fugacity on temperature at constant pressure and composition is given by Eq. (2). Here H_i^* is the enthalpy per mole for the

$$\left(\frac{\partial \ln f_i}{\partial T}\right)_{P,\mathrm{comp}} = \frac{H_i^* - \overline{H}_i}{RT^2} \qquad (2)$$

constituent in the gas phase at very low pressure, and $\overline{H}_i$ its contribution per mole to the enthalpy of the system for the state under consideration. *See* ACTIVITY (THERMODYNAMICS); CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; GAS; PHASE EQUILIBRIUM; SOLUTION.

Paul Bender

Bibliography. K. S. Pitzer and L. Brewer, *Thermodynamics*, rev. ed., 1961; C. E. Reid, *Chemical Thermodynamics*, 1990; P. A. Rock, *Chemical Thermodynamics*, 1983.

Fullerene

A hollow, pure carbon molecule in which the atoms lie at the vertices of a polyhedron with 12 pentagonal faces and any number (other than one) of hexagonal faces. The fullerenes were discovered as a consequence of astrophysically motivated chemical physics experiments that were interpreted by using geodesic architectural concepts. Fullerene chemistry, a field that appears to hold much promise for

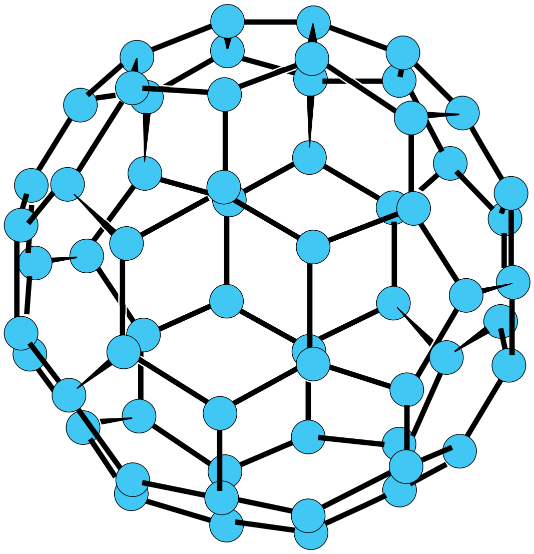


Fig. 1. Structure of fullerene-60 (C_{60}).

materials development and other applied areas, was born from pure fundamental science.

Buckminsterfullerene (C_{60} or fullerene-60) (**Fig. 1**) is the archetypal member of the fullerenes. Other stable members of the fullerene family have similar structures (**Fig. 2**). The fullerenes can be considered, after graphite and diamond, to be the third well-defined allotrope of carbon. Macroscopic amounts of various fullerenes were first isolated in 1990, and since that time it has been discovered that members of this class of spheroidal organic molecules have numerous novel physical and chemical properties. The fullerenes promise to have synthetic, pharmaceutical, and industrial applications. Derivatives have been found to exhibit fascinating electrical and magnetic behavior, in particular superconductivity and ferromagnetism. *See* CARBON; DIAMOND.

Structures. In the fullerene molecule an even number of carbon atoms are arrayed over the surface of a closed hollow cage. Each atom is trigonally linked to its three near neighbors by bonds that delineate a polyhedral network, consisting of 12 pentagons and *n* hexagons. (Such structures conform to Euler's theorem for polyhedrons in that *n* may be any number other than one, including zero.) All 60 atoms in fullerene-60 are equivalent and lie on the surface of a sphere distributed with the symmetry of a truncated icosahedron. The 12 pentagons are isolated and interspersed symmetrically among 20 linked hexagons; that is, the symmetry is that of a modern soccerball. The molecule was named after R. Buckminster Fuller, the inventor of geodesic domes, which conform to the same underlying structural formula. Three of the four valence electrons of each carbon atom are involved in the sp^2 sigma-bonding skeleton, and the fourth *p* electron is one of the 60 involved in a pi-delocalized molecular-orbital electron sea that covers the outside (exo) and inside (endo) surface of the molecule. The resulting cloud of pi-electron density is similar to that which covers the surface of graphite; indeed, the molecule can be considered a round form of graphite. The smallest member of the family, fullerene-20, is dodecahedrene, whose fully hydrogenated analog ($C_{20}H_{20}$) has been created by rational synthesis (production of a compound by using a sequence of chemical reaction steps strategically chosen for the purpose). Several members of the family have been isolated, in particular C_{70}, C_{76}, C_{78}, and C_{84}. Giant fullerenes with at least 600 atoms also seem to form, and appear to have interesting polyhedral shapes. Some species have highly symmetric icosahedral structures, such as C_{240} and C_{540} (**Fig. 3**). *See* ELECTRON CONFIGURATION; GRAPHITE.

General properties. In benzene solution, fullerene-60 is magenta and fullerene-70 red. Fullerene-60 forms translucent magenta face-centered cubic (fcc) crystals that sublime. The ionization energy is 7.61 eV and the electron affinity is 2.6–2.8 eV. The strongest absorption bands lie at 213, 257, and 329 nanometers. Studies with nuclear magnetic resonance spectroscopy yield a chemical shift of 142.7 ppm; this result is commensurate with an aromatic system.

Solid C_{60} exhibits interesting dynamic behavior in that at room temperature the individual round molecules in the face-centered cubic (fcc) crystals are rotating isotropically (that is, freely) at around 10^8 Hz. At around 260 K (8.3°F) there is a phase transition to a simple cubic (sc) lattice accompanied by an abrupt lattice contraction. Rotation is no longer free, and the individual molecules make rotational jumps between two favored (relative) orientational configurations—in the lower-energy one a double bond lies over a pentagon, and in the other it lies over a hexagon. At 90 K (−300°F) the individual molecules stop rotating altogether, freezing into an orientationally disordered crystal involving a mix of the two configurations. *See* CRYSTAL STRUCTURE.

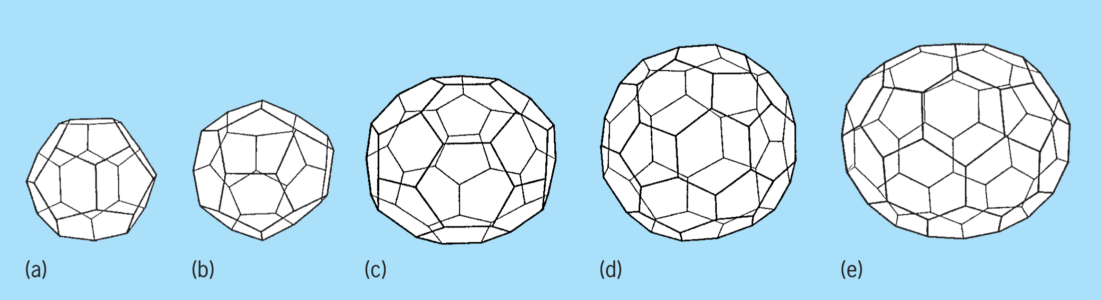


Fig. 2. Some of the more stable members of the fullerene family. (*a*) C_{28}. (*b*) C_{32}. (*c*) C_{50}. (*d*) C_{60}. (*e*) C_{70}.

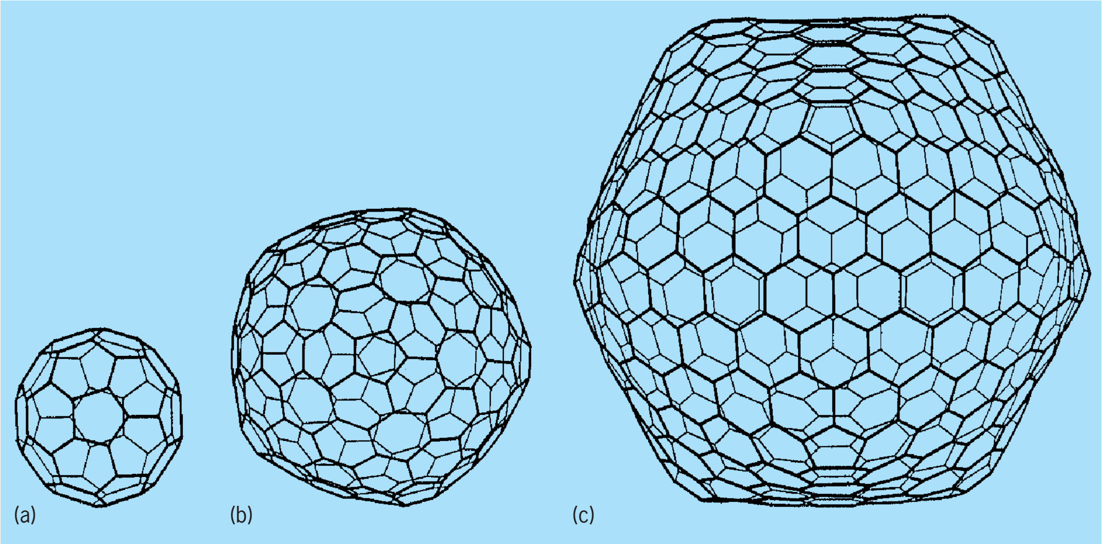


Fig. 3. **Comparison of the structure of fullerene-60 with that of two giant fullerenes. (*a*) C_{60}. (*b*) C_{240}. (*c*) C_{540}.**

Chemistry. Fullerene-60 behaves as a soft electrophile, a molecule that readily accepts electrons during a primary reaction step. It can accept three electrons readily and perhaps even more. The molecule can be multiply hydrogenated, methylated, ammonated, and fluorinated. It forms exohedral complexes in which an atom (or group) is attached to the outside of the cage, as well as endohedral complexes in which an atom is trapped inside the cage. Evidence for oxygen addition compounds, such as $C_{60}O$ and $C_{70}O$, has also been reported. Fullerene-60 appears to catalyze the formation of singlet oxygen in the presence of light. This ability may be related to the observation that under certain conditions solutions of fullerene-60 decompose in the presence of oxygen and light.

Perhaps the most significant chemical result to have been revealed since macroscopic amounts of fullerenes became available in 1990 is that, for synthetic purposes, the C_{60} molecule behaves as though it has only a single resonance form—one in which the 30 double bonds are localized in the bonds that interconnect the pentagons. This is a key factor, as addition to these double bonds is the most important reaction as far as the application of C_{60} in synthesis is concerned.

C_{60} (and other fullerenes) behaves like an electrophilic polyene, and it is susceptible to relatively facile addition by nucleophiles. The pattern of reactivity of the surface network and the way in which specific fullerene derivatives may be produced are gradually unfolding. Addition derivatives with halogens, hydroxyl, phenyl, cyclopentadienel, sugars, and other moieties have been produced:$C_{60}Cl_6$, $C_{60}Br_6$ (**Fig. 4**), $C_{60}Br_8$, $C_{60}Br_{24}$, $C_{60}F_{36}$, $C_{60}Ph_6$, $C_{60}Ph_8$, $C_{60}HPh_5$, $C_{60}(CpH)_6$, and so on. Free-radical additions to produce stable $C_{60}R_n\cdot$ radicals ($R = CH_3$ and $n = 1$–5, and so on) are particularly facile. Transition-metal complexes such as $C_{60}O_2OsO_2$ (4-*t*-butylpyridine)$_2$, $C_{60}Ir(CO)Cl(PPh_3)_2$, and $C_{60}[Pt(PEt_3)_2]_6$ have been isolated. Host-guest crystal complexes form with cyclodextrin, ferrocene, hydroquinone, and sulfur.

Some progress has been made on the production of endohedral (inside cage) complexes; in particular La@C_{82} and U@C_{28} appear to be stable. It is customary with endohedral complexes in which an atom or molecule is trapped inside a fullerene cage to use nomenclature such as M@C_{60}, where @ signifies that M is inside the cage—in this case C_{60}. Helium (He) atom may, under very high pressure, be forced inside the C_{60} cage to form He@C_{60}. The He atom can be expelled on heating. Helium can also be injected into the cage by high-velocity C_{60}/He beam collision experiments.

The possible hazards associated with working with fullerene-60 have not been determined. However, common sense and the ability to catalyze

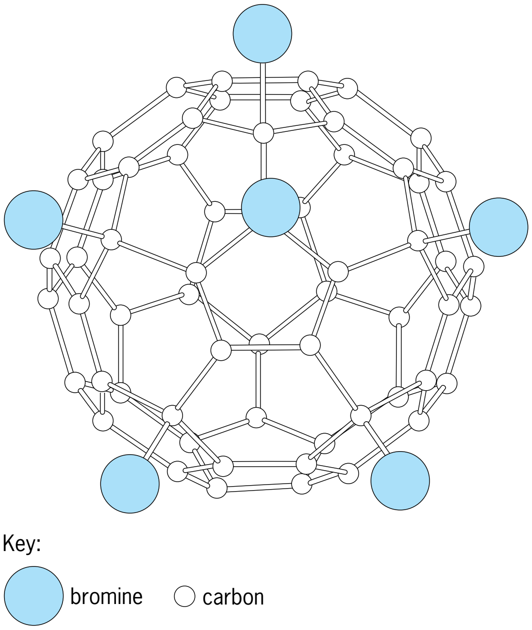


Fig. 4. **Structure of the molecule $C_{60}Br_6$.**

formation of singlet molecules suggest that some care should be exercised when working with it until more is known about possible toxicity.

Superconductivity. On exposure of C_{60} to certain alkali and alkaline earth metals, exohedrally doped crystalline materials are produced that exhibit superconductivity at relatively high temperatures (10 to 33 K or −440 to −400°F). The first to be discovered was the ionic material $K_3{}^+C_{60}{}^{3-}$, which is superconducting at a critical temperature (T_c) of 19 K (−390°F). The highest T_c so far is 33 K (−400°F), exhibited by the mixed intercalate $CsRb_2C_{60}$. Previously, only metallic and ceramic materials exhibited superconductivity at temperatures much greater than a few kelvins. This discovery has opened up the field of superconductivity to a different range of substances—in this case, molecular superconductors. The C_{60} molecule has a triply degenerate lowest unoccupied molecular orbital (LUMO), which in the superconducting materials is half filled, containing three electrons. Other ionic phases, such as M_nC_{60} ($n = 1, 2, 4, 6$, where M is the intercalated metal atom), exist but are not superconducting—they appear to be metallic or semiconductor/insulators. *See* MOLECULAR ORBITAL THEORY; SUPERCONDUCTIVITY.

Nanoparticles and nanotubes. The discovery that graphene networks (single sheets of hexagonally interconnected carbon atoms) can close readily has revolutionized general understanding of certain types of graphitic materials. It has revealed that when isolated aggregates with as many as 10 million carbon atoms anneal, the most stable structure formed is a spheroidal onionlike graphite particle consisting of sets of concentric graphene shells. The discovery of C_{60} itself, together with the nanoparticle studies, clearly shows that dynamic edge effects operating during the graphite formation process govern the resulting intrinsic structures and, in the case of isolated aggregates, result in closed spheroidal structures.

Carbon microparticles can spontaneously rearrange at high temperatures to form onionlike structures in which the concentric shells are fullerene or giant fullerenes. This phenomenon reveals the dynamics of carbon "melting."

Most importantly, carbon nanofibers consisting of concentric graphene tubes (**Fig. 5**) can form. These structures are essentially elongated giant fullerenes and apparently form quite readily. They will be significant in the production of carbon-fiber composite materials.

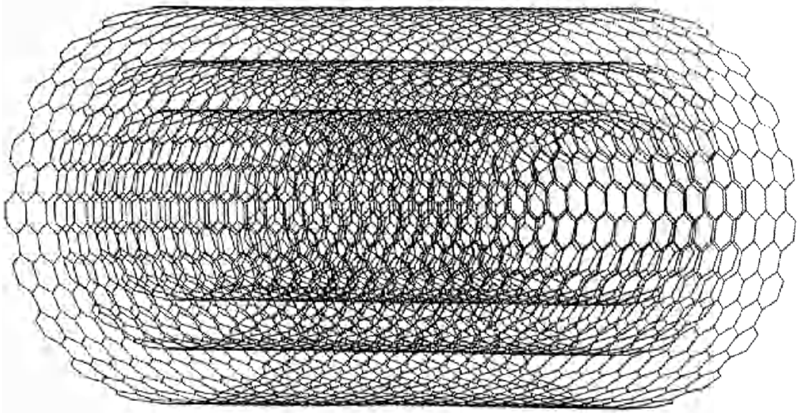


Fig. 5. A section of a carbon nanotube showing the concentric graphene tubes composing it.

Applications. The properties of fullerene materials that have been determined suggest that there is likely to be a wide range of areas in which the fullerenes or their derivatives will have uses. The facility for acceptance and release of electrons suggests a possible role as a charge carrier in batteries. Thin films of C_{60} are transparent at low light intensities and opaque at high intensities, a behavior ideal for optical limiter applications. The fast response means that a thin layer of C_{60} molecules may be used to protect the eye from high transient light pulses. The properties of graphite suggest that lubricative as well as tensile and other mechanical properties of the fullerenes are worthy of investigation. Liquid solutions exhibit excellent properties of optical harmonic generation. The high temperature at which superconducting behavior is observed suggests possible applications in microelectronics devices, as does the detection of ferromagnetism in other fullerene derivatives.

Fullerene formation and implications. Molecular beam experiments indicate that fullerene cages can grow by accretion of carbon, and combined mass spectrometric–synthetic studies have made some progress in developing understanding of the basic fullerene creation process. This process appears to involve the gas-phase aggregation of cumulene/polyyne chains and rings, which then undergo intraaggregate ring cyclizations to produce the closed polycyclic 5/6-ring carbon cage networks.

Perhaps the most important aspect of the fullerene discovery is that the molecule forms spontaneously. This fact has important implications for understanding the way in which extended carbon materials form, and in particular the mechanism of graphite growth and the synthesis of large polycyclic aromatic molecules. It has become clear that as far as pure carbon aggregates of around 60–1000 atoms are concerned, the most stable species are closed-cage fullerenes. A flat, pure graphite carbon sheet has dangling bonds at the edge and these can be satisfied by closure, so that the energy of the system is reduced. In the case of bulk graphite, which consists of stacks of hexagonally arrayed carbon sheets, the interlayer interactions and size effects appear to stabilize the flat sheets. Also, it may be possible to isolate some of the more stable smaller fullerenes, those with fewer than 60 carbon atoms, such as fullerene-50.

Astrophysical implications. Fullerene-60 was discovered as a direct result of physiochemical investigations that simulated processes occurring in stars and in space. Consequently, the likelihood that fullerenes, in particular fullerene-60, and analogs are present in space is a fascinating conjecture. Particularly likely are the exohedral protonated form, $C_{60}H^+$, and exohedral metallated species such as $C_{60}M^+$, where M is an abundant interstellar atom such as sodium, potassium, calcium, or oxygen.

Jonathan Hare; Harold W. Kroto

Bibliography. R. F. Curl and R. E. Smalley, Probing C_{60}, *Science*, 242:1017–1022, 1988; W. Kraetschmer et al., Solid C_{60}: A new form of carbon, *Nature*, 347:354–358, 1990; H. W. Kroto, Space, stars, C_{60}, and soot, *Science*, 242:1139–1145, 1988; H. W. Kroto et al., C_{60}: Buckminsterfullerene, *Nature*, 318:162–163, 1985; H. W. Kroto, W. Allaf, and S. C. Balm, C_{60}: Buckminsterfullerene, *Chem. Rev.*, 91:1213–1235, 1991; K. Prassides and H. W. Kroto, Fullerene physics, *Phys. World*, 5:44–49, 1992; R. Tayor and D. R. M. Walton, The chemistry of fullerenes, *Nature*, 363:685–693, 1993.

Fuller's earth

Any natural earthy material that decolorizes mineral and vegetable oils and has high sorbent capacity for water and oil. The term fuller's earth has no genetic or mineralogic significance. However, the most common earthy materials classed as fuller's earth are calcium montmorillonites and palygorskites (attapulgites) and sepiolites. The term originated in England, where in ancient times raw wool was cleaned by kneading it in water with clay materials that adsorbed dirt and lanolin. The process was known as fulling, and the clay or earth became known as fuller's earth. *See* ADSORPTION; MONTMORILLONITE; SEPIOLITE.

Commercial sources. Several clay deposits in the world are mined and processed for their absorbent, adsorbent, and decolorizing or bleaching properties. Some clays have a high natural decolorizing ability; however, in most instances a clay, normally a calcium montmorillonite, is acid-activated to enhance its bleaching or decolorizing properties. Sulfuric acid is commonly used, and in the treatment process sodium, calcium, magnesium, and iron that occupy the cation exchange sites on the clay surface are removed by the acid and replaced by hydrogen. Also, some aluminum, iron, or magnesium is removed from the mineral structure, increasing the negative charge on the clay surface. These highly charged surfaces covered with hydrogen ions selectively absorb the color bodies and other impurities in the oil. *See* CLAY; CLAY MINERALS.

The largest applications for fuller's earth are as sorbents, and by far the biggest market is pet-litter production. Other large sorbent applications are as carriers for insecticides, pesticides, and fertilizers used in agriculture and as absorbers of oil and water spills on the floors of machine shops, factories, service stations, and other manufacturing plants for safety purposes. All of these sorbent applications use granular particles that are usually of coarse sand size. The particles are dried at high heat to completely drive off absorbed and interlayer water from the clay material and to harden the particles so they do not slake readily in the presence of water or oil.

Occurrence. Fuller's earth is mined in many areas of the world. The most important deposits are located in Georgia, Florida, Mississippi, Tennessee, Illinois, and Missouri in the United States, and in England, Germany, Spain, Russia, Senegal, China, and India.

Haydn H. Murray

Bibliography. H. H. Murray, *Applied Clay Mineralogy Today and Tomorrow*, 1999; R. H. S. Robertson, *Fuller's Earth: A History*, 1986.

Fulminate

A compound containing the —ONC group and derived from fulminic acid, HONC. Fulminates are isomeric with cyanates; that is, cyanates have the same atoms but in different arrangement, —OCN.

The fulminates are commercially important because of the use of mercury fulminate, $Hg(ONC)_2$, in priming compositions and as an initial detonating agent. Mercury fulminate is very sensitive to impact, friction, and heat, but lead azide is replacing mercury fulminate as a detonating agent. *See* AZIDE; CYANATE.

E. Eugene Weaver

Fumigant

A pesticidal chemical or chemical formulation that functions in a gaseous state. Chemical formulations are designed to increase toxicity, reduce flammability, give off warning odors, and provide for sorption at different rates.

Types. Physical types of fumigants include gases, liquids, and solids. Gases, such as methyl bromide and hydrogen cyanide, exist in a gaseous state at normal room temperatures and pressures and must be stored in cylinders under pressure until used. Liquids, such as the 8:20 mixture of carbon tetrachloride and carbon disulfide, volatilize upon exposure to air. Solids, such as aluminum phosphide and calcium cyanide, produce hydrogen phosphide and hydrogen cyanide upon exposure to moisture in air. Most fumigants except hydrogen cyanide are heavier than air.

There are several chemical types of fumigants. These include: halogenated hydrocarbons, such as carbon tetrachloride and ethylene dibromide; sulfur-containing compounds, such as carbon disulfide and sulfur dioxide; cyanides, such as hydrogen cyanide and calcium cyanide; and others, such as phosphine and ethylene dioxide. Although many chemicals are registered as fumigants, extensive use is limited to a few.

Properties. In additon to volatility and toxicity to insects, essential properties of fumigants include diffusion and sorption, all of which affect their use. Heavy fumigant gases diffuse more slowly than lighter gases and do not penetrate as well. All gases diffuse more rapidly at higher temperatures. Heavier-than-air gases diffuse by gravity. The gas molecules may adsorb to the surface of commodities such as grain or they may be absorbed or chemisorbed into the commodities. The extent of sorption is determined by weight and volatility of the gas, the temperature and moisture of the area or commodity, and the presence of extraneous materials such as dust and chaff. Desorption follows at variable rates. An effective fumigant application occurs when toxic concentrations are maintained long enough to be lethal to the pests.

Uses. Fumigants are used in space fumigations to disinfest food-processing plants, warehouses, grain elevators, boxcars, shipholds, stores, and households, and in spot fumigations within those structures. They are used in atmospheric vaults and vacuum chambers and are applied extensively to stacked bags of grain or stored foods under polyethylene sheets, to areas of land to destroy weeds, soil-infesting insects, and nematodes.

Disadvantages. Fumigants have certain inherent disadvantages: Reinfestations may develop as the gas dissipates; off-flavors and odors may result; some are flammable and explosive; some leave toxic residues; certain liquid fumigants are organic solvents; and others are corrosive to metals. All fumigants are toxic to humans; danger exists before, during, and after application. However, fumigants can be applied safely and effectively by trained applicators equipped with safety equipment. For safety reasons, applicators should never work alone.

Action. Fumigant molecules enter the insect's respiratory tubes or penetrate the cuticle and act as respiratory, nerve, or protoplasmic poisons. Some insects close their respiratory tubes after contact with the gas, thus requiring prolonged toxic concentrations. Considerable variability in toxicity exists depending on the fumigant employed and the pest involved. *See* INSECTICIDE; PESTICIDE.

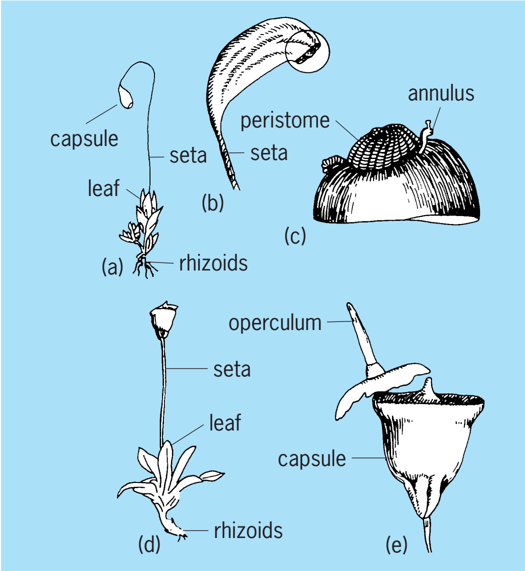
Donald A. Wilbur, Sr.

Bibliography. Institution of Chemical Engineers (England), *A User Guide to Dust and Fume Control*, 1981; H. A. U. Monro, *Manual of Fumigation for Insect Control*, 2d ed., rev. 1969; National Fire Protection Association, *Standards on Fumigation*, 1973.

Funariales

An order of the true mosses (subclass Bryidae). The Funariales, often annuals or biennials and sometimes ephemeral, are for the most part characterized by a uniformity of gametophytic structure in contrast to a variability of sporophytic characters (see **illus.**). Reduction in sporophytic characters is often associated with disturbed habitats and a shortened life cycle.

The plants are generally small or minute, growing scattered or tufted mostly on soil. The stems are erect, and the leaves, inserted in numerous rows, are usually broadly oblong-lingulate and rounded-to-acuminate at the apex, with a single costa. The leaf cells, are lax, thin-walled, and smooth, or less commonly papillose because of projecting upper ends. The sporophytes are terminal (except that in the Funariaceae the sporophyte is formed at the apex of a leading shoot that forms below a basal perigonial bud), with variable seta length. The capsules are generally exserted, erect, and symmetric but sometimes inclined and strongly asymmetric, ovoid to pyriform, and often with a noticeably differentiated sterile neck. The capsules are often inoperculate and lack a peristome, but more often the peristome is present and single or double, the 16 teeth sometimes having endostome segments opposite the teeth and sometimes fused with them. The spores may be spherical or tetrahedral, while the calyptra may be cucullate or mitrate. Chromosome numbers are 9, 11, 12, 13, and 14 or multiples.



Funariales sporophytes. (*a*) *Funaria hygrometrica*, entire plant (*after H. S. Conard, How To Know the Mosses, Jaques, 1944*); (*b*) deoperculate capsule; (*c*) capsule tip enlarged. (*d*) *Physcomitrium turbinatum* (*P. pyriform*), entire plant; (*e*) capsule (*after W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*).

The Funariales include five families, and are heterogeneous in composition. The neck of the capsule is well differentiated, and endostome segments are opposite the peristome teeth. The tendency for stomata to consist of a central slit in a single guard cell gives some unity to the Funariaceae, Gigaspermaceae, and Disceliaceae. The uniformity in gametophytic characters contrasts with a striking complexity and diversity of sporophytic features. The Pseudoditrichaceae have narrowly subulate leaves and long cells, but the segments of the inner peristome are opposite the teeth. The Splachnobryaceae have a peristome consisting essentially of endostome segments and an exostome resembling a very low preperistome (actually a scarcely differentiated primary peristomial cell layer). The Ephemeraceae are minute ephemerals with much reduced gametophytes and sporophytes; the capsules are inoperculate and eperistomate. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA.

Howard Crum

Function generator

An electronic instrument which generates periodic voltage or current waveforms that duplicate various types of well-defined mathematical functions. The simplest function generator usually generates a

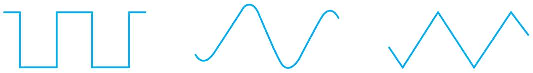


Fig. 1. Waveforms generated by a simple function generator.

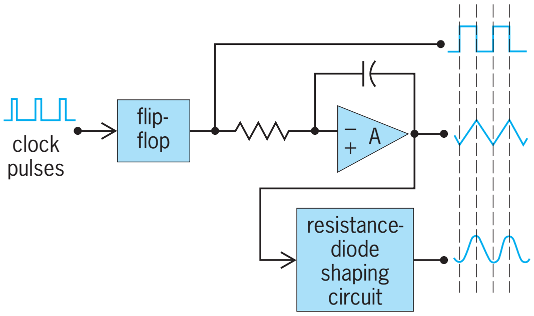


Fig. 2. Elementary function generator.

combination of square waves, triangular waves, and sine waves (**Fig. 1**).

One electronic circuit approach to the design of a simple function generator is to begin with a bistable multivibrator or "flip-flop" controlled in time by a succession of clock pulses which generates the square wave (**Fig. 2**). The triangular waveform is obtained by integrating the square wave through the use of the operational amplifier integrator. The sine wave is obtained by applying the triangular wave to a shaping circuit consisting of a combination of resistors and diodes. *See* AMPLIFIER; MULTIVIBRATOR; WAVE-SHAPING CIRCUITS.

Alternatively the sine wave may be generated by a sinusoidal oscillator (**Fig. 3**). From this output, the square wave may be obtained by amplication, limiting, and clipping of the sine wave. Then the triangular wave may be obtained using an integrator as before. *See* LIMITER CIRCUIT.

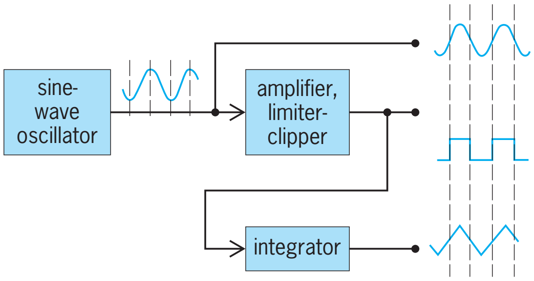


Fig. 3. Alternate form of a function generator.

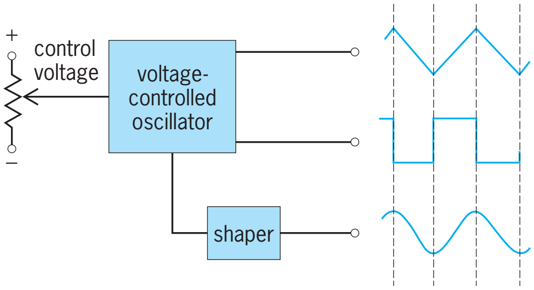


Fig. 4. Voltage-controlled oscillator as a function generator.

A still different configuration of a function generator consists of a circuit, which simultaneously generates triangular waves and square waves, whose rate is controllable by a direct-current level (**Fig. 4**). Such a circuit is identified as a voltage-controlled oscillator (VCO). One of the outputs can be applied to a resistor-diode shaping circuit to obtain the sine wave output. *See* FREQUENCY MODULATOR; PHASE-LOCKED LOOPS.

More sophisticated function generators have circuits built in which will allow many other waveforms to be generated, such as parabolic waveforms and logarithmic functions, using combinations of resistor-diode combinations. Such function generators are widely available as electronic test equipment for general-purpose applications. *See* ELECTRONIC TEST EQUIPMENT.

A combination of counters, programmed read-only memories (PROMs), and a digital-to-analog converter can be used as a function generator, generating almost any function desired to almost any degree of accuracy. *See* COMPUTER STORAGE TECHNOLOGY; DIGITAL-TO-ANALOG CONVERTER. Glenn M. Glasford

Bibliography. H. H. Chiang, *Electronic Wave Forming and Processing Circuits*, 1986, reprint 1989; G. M. Glasford, *Analog Electronic Circuits*, 1986; G. M. Glasford, *Digital Electronic Circuits*, 1988; J. Millman and A. Grabel, *Microelectronics*, 2d ed., 1987; R. B. Northrop, *Analog Electronic Circuits: Analysis and Applications*, 1990.

Functional analysis and modeling (engineering)

The discipline that addresses the activities that a system, software, or an organization must perform to achieve its desired outputs; that is, what transformations are necessary to turn the available inputs into the desired outputs. Additional considerations include the flow of data or items between functions, the processing instructions that are available to guide the transformations, and the control logic that dictates the activation and termination of functions. Functional analysis diagrams have been developed to capture some or all of these concepts.

Functional analysis is performed in systems engineering, software systems engineering, and business process reengineering as a portion of the design process. These design processes typically involve the steps of requirements definition and analysis, functional analysis, physical or resource definition, and operational analysis. This last step of operational analysis involves the marriage of functions with resources to determine if the requirements are met. *See* REENGINEERING; SOFTWARE ENGINEERING; SYSTEMS ENGINEERING.

The concept of examining the logical architecture via functional analysis concurrent with the development of the physical architecture has become a well-accepted principle in the related fields of systems engineering, software engineering, and business process reengineering. Experience in these fields has shown that rushing to a physical representation of a system or organization is likely to produce a set of resources that cannot produce some of the outputs desired by the system. The resulting systems often suffer from gaps in capability, insufficient feedback and control to achieve the objectives desired, and inefficient processing. Functional analysis and modeling is thus a critical activity for the design and analysis of systems and organizations.

Elements. There are four elements to be addressed by any specific functional analysis approach. First, the functions are represented as a hierarchical decomposition, in which there is a top-level function for the system or organization. The top-level function is partitioned into a set of subfunctions that use the same inputs and produce the same outputs as the top-level function. Each of these subfunctions can then be partitioned further, with the decomposition process continuing as often as it is useful.

Second, functional analysis diagrams can represent the flow of data or items among the functions within any portion of the functional decomposition. As the first and subsequent functional decompositions are examined, it is common for one function to produce outputs that are not useful outside the boundaries of the system or organizations. These outputs are needed by other functions in order to produce the needed and expected external outputs.

The **illustration** is a functional analysis diagram for the functional decomposition of an automated teller machine (ATM), showing the preservation of inputs, controls, and outputs. The top-level function, A0, which provides automated bank transactions for customers, is partitioned into three subfunctions: Subfunction A1 services the customer, A2 provides periodic status reports and identifies system problems, and A3 provides access to the ATM for the bank employees and security for the customers.

Processing instructions are a third element that appear in some functional analysis diagrams. These instructions contain the needed information for the functions to transform the inputs to the outputs.

The fourth element is the control flow that sequences the termination and activation of the functions so that the process is both efficient and effective. Sample questions that this information answers are (1) can these functions work serially or must they be processed concurrently, (2) are these functions activated once or a series of times, and (3) what circumstances dictate that one function be activated rather than another function. Included in this control flow are the completion criteria that dictate that a function has finished, and the activation logic that dictates when a function begins.

Approaches. The major functional analysis approaches are structured analysis, data and control flow diagrams, function flow block diagrams and N^2 charts, and behavior diagrams.

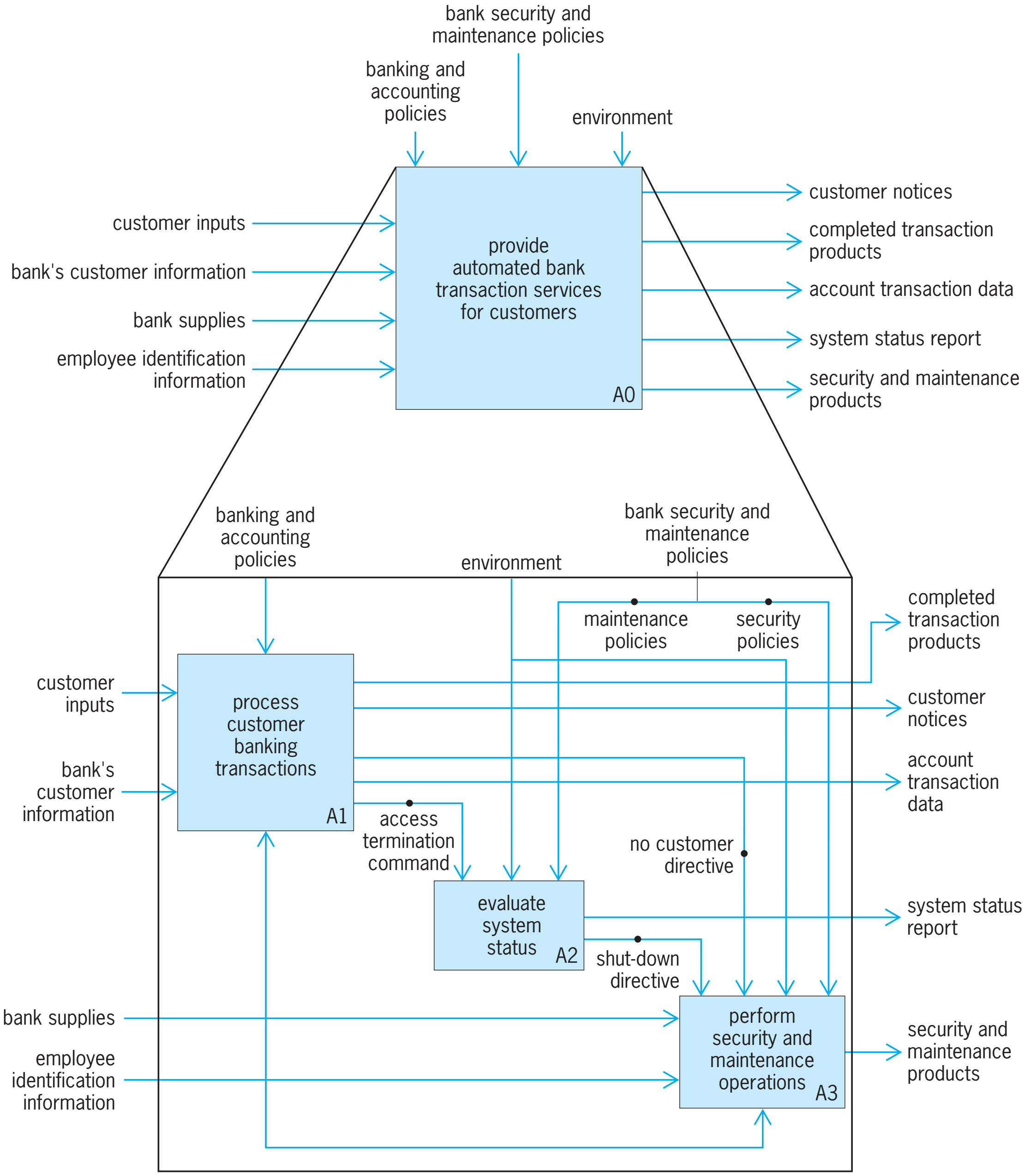
The development of the functional hierarchy can be approached in a top-down (decomposition), bottom-up (composition), or middle-out fashion. Decomposition, often referred to as top-down structuring, begins with the top-level system function and partitions that function into several subfunctions. This decomposition process must conserve all of the inputs to and outputs from the system's top- or zeroth-level function. To conserve inputs or outputs means to use or produce all of them and add no new ones (see illus.). Next, each of the several first-level functions are decomposed (partitioned) into a second level set of subfunctions. Not every function must be decomposed; only those for which additional insight into the production of outputs is needed should be partitioned.

The opposite approach, composition, is a bottom-up approach. With composition, the analyst starts by identifying the simple low-level functions associated with simple scenarios involving only one of the outputs of the system. Each set of low-level functions is a sequence of input, function, output-input, function, output-input, . . . , function, output-input, function, output. These functions usually show up in third, fourth, or even lower levels of the hierarchy. For complex systems this initial step is a substantial amount of work. After the many sets of low-level functions have been defined, the analyst begins the process of grouping the functions into similar groups. These groups are, in turn, aggregated into larger similar groups; this process continues until a hierarchy is formed from bottom to top. The composition method dates back to the 1960s and 1970s when systems engineering was in its infancy; many systems engineers continue to prefer this approach. There is no empirical evidence that either the composition approach or decomposition approach is better.

Ultimately, using a combination of decomposition and composition approaches is wisest. This is sometimes referred to as middle-out. Often, the analyst makes use of a set of low-level functions associated with specific scenarios defined in the operational concept to establish a "sense" of the system, and then posits a top-level decomposition that is likely to meet the needs of the functional analysis.

Feedback and control. Feedback plays an important role in functional analysis and modeling. Feedback and control is the comparison of the actual characteristics of an output with desired characteristics of that output for the purpose of adjusting the process of transforming inputs into that output. Open-loop control processes may or may not make this measurement, but in either case make no adjustments to the process once started. Closed-loop control processes use measurements of the output as feedback for the purpose of adjusting or controlling the transformation process.

Closed-loop control processes contain at least four subprocesses: comparison of current and desired output characteristics, control adjustments to the process based upon the comparison, the transformation process for turning inputs into outputs, and a sensing process for turning the output into measured dimensions that can be compared to the desired



Functional decomposition of an automated teller machine's functions in an IDEF0 model, showing the preservation of inputs, controls, and outputs.

output. The first element is the comparison process in which current values of key variables are compared with desired values of those variables. The comparison process requires definition in advance for what elements of the state of the process are going to be compared. This comparison inevitably introduces a time lag into the process. This element of the feedback process is trivial but at the same time is the cornerstone of the process. The second element is the control process for deciding what to do about the difference between the current value of the output and the desired value of the output. The third element of the feedback process is the transformation process that is being controlled by the feedback process. This process dictates a great deal about how a successful feedback process should be created and is often adapted by the feedback process as part of the correction activity. Sensing the output of the process being controlled is the final element of the feedback process. *See* CONTROL SYSTEMS. Dennis M. Buede

Bibliography. D. M. Buede, *The Engineering Design of Systems*, Wiley, New York, 2000; D. J. Hatley, P. Hruschka, and I. A. Pirbhai, *Process for System Development: A Practical Guide to Requirements and Modeling*, Dorset House, New York, 2000; D. W. Oliver, T. P. Kelliher, and J. P. Keegan, Jr., *Engineering Complex Systems with Models and Objects*, McGraw-Hill, New York, 1997; R. J. Wieringa, Combining static and dynamic modeling methods: A comparison of 4 methods, *Comput. J.*, 38(1):17–30, 1995.

Fundamental constants

The set of physical quantities that play a fundamental role in basic physical theories, primarily at the atomic and subatomic levels, including also those quantities that provide the connection between microscopic physics and the macroscopic world. Examples of fundamental constants are the velocity of light in vacuum c; the Planck constant h; the universal gravitation constant G; the electron charge e; the masses of the proton, the neutron, and the electron, m_p, m_n, and m_e; the Boltzmann constant k; and the Avogadro constant N_A; as well as any physical constant such as the fine-structure constant α that is a combination of these constants.

These quantities typify the scope of the field. The speed of light c and the gravitation constant G provide the scale for Einstein's theory of relativity; Planck's constant h is the fundamental magnitude of quantum mechanics. The fine-structure constant α and the masses of the fundamental particles, together with c and h, are the basic quantities that define the structure of atoms and molecules and provide the foundation for chemistry and atomic physics. The Avogadro constant N_A is the connection between the atomic and macroscopic scales. *See* ATOMIC STRUCTURE AND SPECTRA; AVOGADRO'S NUMBER; PLANCK'S CONSTANT; QUANTUM MECHANICS; RELATIVITY.

As considered here, the rubric of fundamental constants combines the fields of metrology, statistics, and data analysis to determine the best estimates of their values. Why the physical constants have the values that they do will not be discussed. For one thing, that question strictly can be asked only of dimensionless quantities (quantities of dimension 1), such as m_p/m_e, α, or $Gm_p{}^2/hc$, since the value of a quantity with dimensions depends on the arbitrary units used to measure it. As an example, since 1983 the velocity of light c has been defined to be 299 792 458 m/s, thus defining the meter as the distance light travels in a specified time as determined by an ideal clock. "Why is c so much larger than 'ordinary' speeds?" is a different kind of question, and finds its answer in atomic physics and quantum mechanics. *See* LIGHT.

Some fundamental physical quantities can be measured directly in terms of the basic units of measurement, but in most cases an experiment yields only a combination such as m_p/m_e, h/e, N_Ah, or e^2/h. The fine-structure constant α is proportional to e^2/h and may be considered to be a more fundamental constant than e itself. In the unrationalized (three-dimensional) centimeter-gram-second (cgs) system of units, $\alpha = e^2/\hbar c$ (where $\hbar c = h/2\pi$); in the modern rationalized, four-dimensional International System of Units (SI), this becomes $\alpha = [e^2/(4\pi\epsilon_0)]/\hbar c = \mu_0 ce^2/(2h)$, where ϵ_0 is the permittivity of vacuum and $\mu_0 = 1/(\epsilon_0 c^2) \equiv 4\pi \times 10^{-7}$ N/A^2 is the permeability of vacuum.

No relationship, such as the foregoing one for α, is known between the gravitational constant G and other physical constants, and one of the major challenges in theoretical physics is grand unification, which attempts to combine gravitation, electromagnetism, and nuclear forces into a single entity, presumably establishing a relationship between G and α, h, c, and the masses of elementary fundamental particles. *See* GRAND UNIFICATION THEORIES; GRAVITATION.

Importance of measurement. The interest in precise measurements is due to the existence of quantities in nature that are invariable. Without the development of quantum mechanics at the beginning of the twentieth century, there could be no basis for the assumption that the masses of all protons were exactly the same. Whereas the mass of a baseball is easily measured to within a fraction of a percent, the measurement applies only to that one ball, and there is little need to be more precise, since the masses of legal baseballs may vary by as much as ±2.5%. The mass of a proton or an electron, however, is known to within 1 part in 10^7, and because of the quantization of physical quantities this applies to individual particles, not merely to the mean mass of a large number of particles. Since the properties of atoms and molecules are the same wherever and whenever they are measured (and there is no evidence to contradict the assumption that this holds throughout the universe), the atomic constants provide a means for preserving and reproducing standards of measurements. Except for the unit of mass, the world's measurement standards are derived from the speed of light and the properties of atomic systems rather than from arbitrary artifacts. *See* PHYSICAL MEASUREMENT.

A second important reason for determining the values of the atomic constants with all possible precision and accuracy (precision referring to the reproducibility of the result of repeated measurement, and accuracy referring to the degree to which the measurement gives the true or correct value) is that the only test of the validity of physics is the ability to predict the outcome of experiment. If two determinations of a physical property do not agree within the uncertainties of the measurements, either the experimental measurement or the underlying theory of at least one of the determinations is wrong.

An early example of precise measurement leading to new physics was Lord Rayleigh's attempt in 1895 to improve the precision of the determination of the atomic weight of nitrogen. He found that after the oxygen, carbon dioxide, and water vapor had been removed from air the density of the remaining gas (presumed to be pure nitrogen) was greater by 0.5% than that of nitrogen prepared from chemical sources such as ammonia. This observation, and the conviction that the measurements were sufficiently precise that the discrepancy could not be the result of experimental error, led to the identification of the previously unknown gas argon and ultimately to a whole new family of noble gases.

Another example is the discovery of the Lamb shift. According to the Dirac theory of atomic spectra, the S and $P_{1/2}$ states in atomic hydrogen should have the same energy, but in 1947 W. Lamb, Jr., and

R. C. Retherford, in a measurement of microwave absorption spectra, showed that the $2S$ level was 1.3 parts per million (ppm) higher than the $2P_{1/2}$ level. The search for the explanation of this shift led to the development of quantum electrodynamics (QED). Indications of such a shift had actually been seen in the optical spectrum as much as 20 years earlier, but the precision of those measurements was sufficiently low that it was not generally recognized as a real effect. *See* QUANTUM ELECTRODYNAMICS.

History of measurements. The attempts to determine the numerical values of physical constants date back to the measurements of c and G in the seventeenth and eighteenth centuries. Following the development of the Bohr theory of the atom and quantum mechanics in the first three decades of the twentieth century, precise measurements were made of the elementary charge e and the ratios e/h and e/m_e. Technical developments in spectroscopy led to successively more accurate determinations of the Rydberg constant R_∞, while x-ray spectroscopy and crystallography led to determinations of the Avogadro constant N_A. Microwave technology developed during and after World War II greatly broadened the capability for precise measurement and made possible microwave spectroscopy of the fine structure and hyperfine structure of hydrogen which, with a sufficiently refined theory, provide values of the fine-structure constant α. In addition, nuclear magnetic resonance techniques yielded measurements of the ratios of electron and proton magnetic moments and cyclotron frequencies, which give m_p/m_e. The proton gyromagnetic moment γ_p combined with the proton magnetic moment ratio μ_p/μ_N provides a determination of the Faraday constant, given in Eq. (1), that is free of

$$F = N_A e = \frac{\gamma_p M_p}{\mu_p/\mu_N} \qquad (1)$$

the problems associated with the classical electrochemical determination of that quantity. (Here μ_p is the proton magnetic moment, μ_N is the nuclear magneton, and M_p is the proton molar mass.) *See* ELECTROCHEMICAL EQUIVALENT; ELECTRON; ELECTRON SPIN; FINE STRUCTURE (SPECTRAL LINES); GYROMAGNETIC RATIO; HYPERFINE STRUCTURE; MAGNETISM; MAGNETON; MICROWAVE SPECTROSCOPY; NUCLEAR MAGNETIC RESONANCE (NMR); RYDBERG CONSTANT.

The study of atomic constants received an unexpected stimulus from solid-state physics in 1967. The Josephson effect in superconductors provided the basis for a determination of the ratio $2e/h = K_J = nf/V(n)$, where f is the frequency of the microwave field on the junction and $V(n)$ is the potential difference of the nth step. Until this discovery, and the surprising demonstration of the precision achievable in the measurement of this ratio, the determination of e/h was based on measurements of the photoelectric effect at optical wavelengths or its equivalent, the short-wavelength limit of the continuous x-ray spectrum. These measurements were inherently much less precise than the Josephson effect, which determined the voltage-wavelength product (or the voltage/frequency ratio) in the microwave region of the electromagnetic spectrum. The value of $2e/h$ could be combined with the proton gyromagnetic ratio to determine a value of the fine-structure constant that was independent of atomic spectroscopy and of any uncertainties of quantum electrodynamics. *See* JOSEPHSON EFFECT.

An analysis of the atomic constants in 1963 had to deal with two incompatible values of the fine-structure constant determined from the microwave spectrum of hydrogen, one of which included a contribution dependent on the structure of the proton. If the proton is not an elementary particle but has an internal structure, its center of charge in a hydrogen atom would be affected by the electromagnetic field of the orbiting electron and might not be identical with its center of mass. This recoil effect had been estimated to produce a shift in the hyperfine interval that might be as large as 35 ppm. This uncertainty of the contribution required that the measurement be excluded from the analysis. However, by 1969 additional theoretical studies had indicated that these corrections were in fact no more than 2–3 ppm. The value of α deduced from the Josephson effect and the proton gyromagnetic ratio was then in excellent agreement with the hyperfine-structure value. The fine-structure value of α was thus shown to be too low by 20 ppm. More significantly, since the solid-state value α was independent of quantum electrodynamics, the validity of that theory could now be tested unambiguously by comparing theory and experiment without using those experiments at the same time to determine α. *See* PROTON.

This is an example of the unity of science: a low-temperature solid-state measurement gave information that resolved a problem in atomic physics (fine-structure spectrum of hydrogen, characteristic energy ~0.1 meV) and verified a calculation associated with an important question in high-energy physics (proton structure, characteristic energy >10 GeV).

Another connection between solid-state and atomic physics, discovered by K. von Klitzing in 1980, is the quantization of the ratio of the transverse potential (V) to the longitudinal current (I; Hall coefficient) in a two-dimensional electron gas at a semiconductor heterojunction at high magnetic field and sufficiently low temperature.

Specifically, when the longitudinal potential approaches zero, this ratio is given by Eq. (2), where

$$\frac{V}{I} = \frac{R_K}{i} = \frac{h}{e^2 i} \qquad (2)$$

i is an integer. The von Klitzing effect provides a determination of the fine-structure constant as well as a reproducible and drift-free resistance standard that is much more convenient than that provided by the calculable capacitor. *See* HALL EFFECT.

The level of precision of measurements has improved steadily. In one of the earliest careful studies in this field, by R. T. Birge in 1929, the uncertainty of the atomic constants was typically 1 part in 500 to 1 part in 1000 (except for c with a precision of 20 ppm and R_∞ with 0.15 ppm). By 1950 the level of precision of the constants had improved by a factor of roughly 50, to the order of 30 ppm. In the 1998 analysis the level of uncertainty had generally been reduced to levels of 0.1 ppm to 0.01 ppm. Thus, over a period of 70 years the precision of the determinations has improved by a factor of 5 every 10 years.

Atomic standards for electrical units. The improvements in precision of the fundamental constants have had an important influence on electrical engineering. Units for electrical quantities are defined from the laws of electromagnetism (Ampère's law, Biot-Savart law, and Ohm's law) and the requirement for coherence between the units of mechanical and electrical energy. As a practical matter, however, the units of electrical potential and electrical resistance had been maintained by using electrolytic cells and resistance coils, respectively. These artifact standards must be calibrated in terms of the SI definitions, but when the measurements of the fundamental constants had a relative precision of 10^{-4} to 10^{-3}, the uncertainty in the consistency of the laboratory electrical standards with their definitions was unimportant. When the relative precision improved to the level of 10^{-5} and better, these calibrations had to be included with the other pertinent calibration quantities in the analysis of the values of physical constants. *See* AMPÈRE'S LAW; BIOT-SAVART LAW; OHM'S LAW.

Prior to 1950, electrical units could be evaluated only by using a current balance to compare (at a level of the order of 1–10 ppm) electrical and mechanical forces on a current-carrying coil. The development of the Thompson-Lampard calculable capacitor made it possible to calibrate a laboratory standard of resistance directly, depending only on geometrical measurements and the value of c, while the Josephson effect provided a means to monitor the constancy of voltage standards. Beginning in 1973, many national laboratories adopted a representation of the volt using the definition $(2e/h)V_{\text{LAB}} = 483{,}594$ GHz, where V_{LAB} is the volt standard. (The United States had used the frequency 483,593.42 GHz since July 1972 to maintain its standard of the volt.) It then became evident, based on such calibrations and on international intercomparisons, that resistance coils and standard cells can drift in value by parts per million per decade. *See* CAPACITANCE MEASUREMENT.

Based on the results of the 1986 adjustment of the physical constants and on subsequent additional measurements of both the Josephson effect and the von Klitzing effect, new standards for the volt ($2e/h = 493{,}597.9$ GHz/V_{90}) and the ohm ($R_H = 25{,}812.807\ \Omega_{90}$) became effective on January 1, 1990. A realization of the volt in terms of the Josephson effect reproduces the SI volt with an

Recommended values (2002) of selected fundamental physical constants

Quantity	Symbol	Numerical value[*]	Units[†]	Relative uncertainty (standard deviation)
Speed of light in vacuum	c	299792458	m/s	(defined)
Permeability of vacuum	μ_0	$4\pi \times 10^{-7}$	N/A^2	(defined)
Permittivity of vacuum	ϵ_0	8.854187817. . .	10^{-12} F/m	(defined)
Constant of gravitation	G	6.6742 (10)	10^{-11} m^3/(kg · s^2)	1.5×10^{-4}
Planck constant	h	6.6260693 (11)	10^{-34} J · s	1.7×10^{-7}
Elementary charge	e	1.60217653 (14)	10^{-19} C	8.5×10^{-8}
Magnetic flux quantum, $h/(2e)$	Φ_0	2.06783372 (18)	10^{-15} Wb	8.5×10^{-8}
Fine-structure constant,	α	7.297352568 (24)	10^{-3}	3.3×10^{-9}
$\mu_0 ce^2/(2h)$	α^{-1}	137.03599911 (46)		3.3×10^{-9}
Electron mass	m_e	9.1093826 (16)	10^{-31} kg	1.7×10^{-7}
Proton mass	m_p	1.67262171 (29)	10^{-27} kg	1.7×10^{-7}
Neutron mass	m_n	1.67492728 (29)	10^{-27} kg	1.7×10^{-7}
Proton-electron mass ratio	m_p/m_e	1836.15267261 (85)		4.6×10^{-10}
Rydberg constant, $m_e c\alpha^2/(2h)$	R_∞	10973731.568525 (73)	m^{-1}	6.6×10^{-12}
Bohr radius, $\alpha/(4\pi R_\infty)$	a_0	5.291772108 (18)	10^{-11} m	3.3×10^{-9}
Compton wavelength of the electron, $h/(m_e c) = \alpha^2/(2R_\infty)$	λ_c	2.426310238 (16)	10^{-12} m	6.7×10^{-9}
Classical electron radius, $\mu_0 e^2/(4\pi m_e) = \alpha^3/(4\pi R_\infty)$	r_e	2.817940325 (28)	10^{-15} m	1.0×10^{-8}
Bohr magneton, $eh/(4\pi m_e)$	μ_B	9.27400949 (80)	10^{-24} J/T	8.6×10^{-8}
Electron magnetic moment	μ_e	−9.28476412 (80)	10^{-24} J/T	8.6×10^{-8}
Electron magnetic moment/Bohr magneton ratio	μ_e/μ_B	−1.0011596521859 (38)		3.8×10^{-12}
Nuclear magneton, $eh/(4\pi m_p)$	μ_N	5.05078343 (43)	10^{-27} J/T	8.6×10^{-8}
Proton magnetic moment/nuclear magneton ratio	μ_p/μ_N	2.792847351 (28)		1.0×10^{-8}
Avogadro constant	N_A	6.0221415 (10)	10^{23}	1.7×10^{-7}
Faraday constant, $N_A e$	F	96485.3383 (83)	C/mol	8.6×10^{-8}
Molar gas constant	R	8.314472 (15)	J/(mol · K)	1.7×10^{-6}
Boltzmann constant, R/N_A	k	1.3806505 (24)	10^{-23} J/K	1.8×10^{-6}

[*] As recommended by the CODATA Task Group on Fundamental Constants. The digits in parentheses represent the one-standard-deviation uncertainties in the final two digits of the quoted value.

[†] A = ampere, C = coulomb, F = farad, J = joule, kg = kilogram, K = kelvin, m = meter, mol = mole, N = newton, s = second, T = tesla, Wb = weber.

overall relative uncertainty of 4 parts in 10^7 (one standard deviation), while a realization of the ohm in terms of the von Klitzing effect reproduces the ohm with an uncertainty of 2 parts in 10^7. The watt balance measures the force on a current-carrying coil in a constant magnetic field. This force is equal to BIl, where B is the magnetic field, I is the current, and l is the length of the coil, and it is balanced by a gravitational force mg, where m is the mass of the coil and g is the acceleration of gravity. Then, with the current source removed, the coil is moved at constant speed v in the same field, and the induced potential difference $V = Blv$ between its ends is measured. Hence it is possible to equate the electrical and mechanical power, $VI = mgv$. Since the electrical units are maintained in term of $R_K = h/e^2$ and $K_J = 2e/h$, these measurements allow a determination of the Planck constant h. Such measurements not only provide a means to monitor possible changes in the mass of the International Prototype of the kilogram but can ultimately lead to its redefinition in terms of invariant atomic standards. *See* ELECTRICAL UNITS AND STANDARDS; RESISTANCE MEASUREMENT; VOLTAGE MEASUREMENT; WATT BALANCE.

Analysis of experimental data. The various experimental measurements of the atomic constants are not independent. For example, the fine-structure constant may be determined from atomic spectroscopy of hydrogen; from the quantum Hall resistance (von Klitzing effect); by combining R_∞, the Josephson constant $2e/h$, the proton gyromagnetic moment, and the ratio of the proton nuclear magnetic resonance frequency to the electron cyclotron frequency; or from a measurement of the anomalous moment of the electron (coupled with a calculation of its theoretical expression). Similarly, the Avogadro constant can be determined directly by measuring the density of a crystal, its atomic lattice spacings, and its molar mass, and indirectly by combining measurements of the von Klitzing constant, the Josephson constant, and the Faraday constant; while the Faraday constant, measured directly by the electrolysis of silver, is also indirectly deduced from measurements of the proton gyromagnetic moment, using Eq. (1).

Because of such relationships among the measurements of the atomic constants, all of the data, including the calibration factors relating the international standards of the ohm and the volt to their SI definitions, must be analyzed as a single entity. In determining a "best" value of a single quantity from repeated measurements, the average is generally used, often with statistical weights so that the more reliable data are given more importance in the average. The formalism of least squares is an appropriate generalization of this process when more than one quantity is to be determined from a set of data. *See* LEAST-SQUARES METHOD.

Since 1969 the adjustment of the fundamental constants (see **table**) has been carried out under the auspices of the CODATA Task Group on Fundamental Constants. CODATA, the Committee on Data for Science and Technology, is an interdisciplinary committee of the International Council of Scientific Unions.

E. Richard Cohen

Bibliography. G. Cohen-Tannoudji, *Universal Constants in Physics*, Horizons of Science Series, McGraw-Hill, New York, 1992; W. H. McCrea and M. J. Ree (eds.), *The Constants of Physics* (Proceedings of a Royal Society discussion meeting held May 25–26, 1983, London), 1985; P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 1998, *Rev. Mod. Phys.*, 72(2):351–495, 2000; P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 2002, *Rev. Mod. Phys.*, 77(1):1–107, 2005; B. W. Petley, *The Fundamental Physical Constants and the Frontier of Measurement*, Adam Hilger, Boston, 1988.

Fundamental interactions

Fundamental forces that act between elementary particles, of which all matter is assumed to be composed.

Properties of Interactions

At present, four fundamental interactions are distinguished. The properties of each are summarized in the **table**.

Gravitational interaction. This interaction manifests itself as a long-range force of attraction between all elementary particles. The force law between two particles of masses m_1 and m_2 separated by a distance r is well approximated by the newtonian expression $G_N(m_1m_2/r^2)$, where G_N is the newtonian constant, equal to $6.6742 \pm 0.0010 \times 10^{-11}\ \mathrm{m^3 \cdot kg^{-1} \cdot s^{-2}}$. The dimensionless quantity $(G_N m_e m_p)/(\hbar c)$ is usually taken as the constant characterizing the gravitational interaction, where m_e, and m_p are the electron and proton masses, $2\pi\hbar$ is Planck's constant, and c is the velocity of light. *See* GRAVITATION.

Electromagnetic interaction. This interaction is responsible for the long-range force of repulsion of like, and attraction of unlike, electric charges. The dimensionless quantity characterizing the strength of electromagnetic interaction is the fine-structure constant, given by Eq. (1) in SI units, where e is the electron charge and ϵ_0 is the permittivity of empty space.

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{(137.03599911 \pm 0.00000046)} \qquad (1)$$

At comparable distances, the ratio of gravitational to electromagnetic interactions (as determined by the

Properties of the four fundamental interactions

Interaction	Range	Exchanged quanta
Gravitational	Long-range	Gravitons (*g*)
Electromagnetic	Long-range	Photons (γ)
Weak nuclear	Short-range ≈ 10^{-18} m	W^+, Z^0, W^-
Strong nuclear	Short-range ≈ 10^{-15} m	Gluons (*G*)

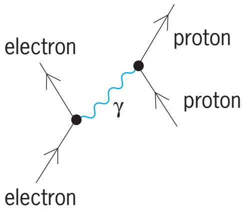


Fig. 1. Feynman diagram of electromagnetic interaction between an electron and a proton.

strength of respective forces between an electron and a proton) is given by the quantity $4\pi\epsilon_0 G_N m_e m_p/e^2$, which is approximately 4×10^{-37}. *See* COULOMB'S LAW; ELECTROSTATICS.

In modern quantum field theory, the electromagnetic interaction and the forces of attraction or repulsion between charged particles are pictured as arising secondarily as a consequence of the primary process of emission of one or more photons (particles or quanta of light) emitted by an accelerating electric charge (in accordance with Maxwell's equations) and the subsequent reabsorption of these quanta by a second charged particle. The space-time diagram (introduced by R. P. Feynman) for one photon exchange is shown in **Fig. 1**. A similar picture may also be valid for the gravitational interaction (in accordance with the quantum version of A. Einstein's gravitational equations), but with exchanges of zero-rest-mass gravitons (g) rather than zero-rest-mass photons. (The existence of the graviton, however, has not yet been experimentally demonstrated.) *See* FEYNMAN DIAGRAM.

In accordance with this picture, the electromagnetic interaction (to one photon exchange approximation) is usually represented by reaction (2), where

$$e + P \rightarrow (e + \gamma) + P \rightarrow e + (P + \gamma) \rightarrow e + P \quad (2)$$

γ is the photon, emitted by the electron and reabsorbed by the proton. For this interaction, and also for the gravitational interaction represented by reaction (3), the nature of the participating parti-

$$e + P \rightarrow (e + g) + P \rightarrow e + (P + g) \rightarrow e + P \quad (3)$$

cles (electron e and proton P) is the same, before and after the interaction, and the exchanged quanta (γ or g) are electrically neutral. *See* GRAVITON; LIGHT; MAXWELL'S EQUATIONS; PHOTON; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM GRAVITATION; QUANTUM MECHANICS.

Weak nuclear interactions. The third fundamental interaction is the weak nuclear interaction, which is responsible for the decay of a neutron into a proton, an electron, and an antineutrino. Its characteristic strength for low-energy phenomena is measured by the Fermi constant G_F, which is equal to $1.026 \times 10^{-5}\ m_p{}^{-2}\hbar^3/c$. Unlike electromagnetism and gravitation, weak interactions are short-range, with a force law of the type $e^{-M_W cr/\hbar}$, the range of the force ($\hbar/M_W c$) being of the order of 10^{-18} m. Until 1973, the only known weak interactions were those which changed the nature of the interacting particles (unlike electromagnetism and gravity). For example, consider reactions (4), where P is the proton, N is the

$$P + e^- \xrightarrow{\text{weak}} N + \nu_e \quad \text{(this reaction is equivalent to } \beta \text{ decay of the neutron } N \rightarrow P + e^- + \overline{\nu}_e) \qquad (4a)$$

$$P + \mu^- \xrightarrow{\text{weak}} N + \nu_\mu \quad \text{(muon capture by a proton with the emission of a neutrino)} \qquad (4b)$$

$$\mu^- + \nu_e \xrightarrow{\text{weak}} \nu_\mu + e^- \quad \text{(this reaction is equivalent to muon decay } \mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu) \qquad (4c)$$

neutron, μ^- is the negatively charged muon, ν_e and ν_μ are the electronic and muonic neutrinos, and $\hbar_e$ and $\hbar_\mu$ are the corresponding antineutrinos. In reaction (4*a*), the weak interaction transforms a proton into a neutron and at the same time an electron into a neutrino.

An important question was finally answered in 1983: Is the weak interaction similar to electromagnetism in being mediated primarily by intermediate objects, the W^+ and W^- particles. If this is the case, then reactions (4*a*) and (4*c*), for example, would in detail be represented as reactions (5*a*) and (5*b*).

$$P^+ + e^- \rightarrow (N^0 + W^+) + e^- \rightarrow N^0 + (W^+ + e^-) \rightarrow N^0 + \nu^0 \quad (5a)$$

$$\mu^- + \nu_e^0 \rightarrow (\nu_\mu^0 + W^-) + \nu_e^0 \rightarrow \nu_\mu^0 + (W^- + \nu_e^-) \rightarrow \nu_\mu^0 + e^- \quad (5b)$$

[The superscript on each particle gives its electrical charge (+, 0, −) in units of the proton's charge.] The experimental answer (discovered at the CERN laboratory at Geneva) is that W^+ and W^- do exist, with a mass m_W of 80.9 ± 2.0 GeV/c^2. Each carries a spin of magnitude $\hbar$ just as does the photon (γ). The mass of these particles gives the range [$\hbar/(m_W c) = 10^{-18}$ m] of the weak interaction, and is also related to its strength G_F, as discussed below. *See* INTERMEDIATE VECTOR BOSON.

Another crucial discovery in weak interaction physics was the neutral current phenomenon in 1973, that is, the discovery of new types of weak interactions where (as in the case of electromagnetism or gravity) the nature of the interacting particles is not changed during the interaction, as in reactions (6).

$$\nu_\mu + e^- \xrightarrow{\text{weak}} \nu + e^- \quad (6a)$$

$$\nu_\mu + P \xrightarrow{\text{weak}} \nu_\mu + P \quad (6b)$$

$$\nu_e + N \xrightarrow{\text{weak}} \nu_e + N \quad (6c)$$

$$e^- + P \xrightarrow{\text{weak}} e^- + P \quad (6d)$$

See NEUTRAL CURRENTS.

The 1983 experiments at CERN also gave evidence for the existence of an intermediate particle Z^0 which is believed to mediate such reactions. Thus reaction (6*a*), expressed in detail, is reaction (7).

$$\nu_\mu + e^- \rightarrow (\nu_\mu + Z^0) + e^- \rightarrow \nu_\mu + (Z^0 + e^-) \rightarrow \nu_\mu + e^- \quad (7)$$

The mass m_Z of the Z^0 has been found to be 93.0 ± 2.0 GeV/c^2. The magnitudes of the W^+, W^-, and Z^0 masses had been predicted by the unified theory of electromagnetic and weak interactions (the electroweak interaction, discussed below), 16 years before the experiments which discovered them.

In contrast to gravitation, electromagnetism, and strong nuclear interactions, weak interactions violate left-right and particle-antiparticle symmetries. *See* PARITY (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS); WEAK NUCLEAR INTERACTIONS.

Strong nuclear interaction. The fourth fundamental interaction is the strong nuclear interaction between protons and neutrons, which resembles the weak nuclear interaction in being short-range, although the range is of the order of 10^{-15} m rather than 10^{-18} m. Within this range of distances the strong force overshadows all other forces between protons and neutrons, with a characteristic strength parameter of the order of unity (compared with the electromagnetic strength parameter $\alpha \approx 1/137$).

Protons and neutrons are themselves believed to be made up of yet more fundamental entities, the up (u) and down (d) quarks ($P = uud$, $N = udd$). Each quark is assumed to be endowed with one of three color quantum numbers [conventionally labeled red (r), yellow (y), and blue (b)]. The strong nuclear force can be pictured as ultimately arising through an exchange of zero rest-mass color-carrying quanta of spin $\hbar$ called gluons (G) [analogous to photons in electromagnetism], which are exchanged between quarks (contained inside protons and neutrons), as in reaction (8). Since neutrinos, electrons, and muons (the so-called leptons) do not contain quarks, their interactions among themselves or with protons and neutrinos do not exhibit the strong nuclear force. There is indirect experimental evidence for the existence of the gluons and of their spin being $\hbar$. *See* COLOR (QUANTUM MECHANICS); GLUONS; LEPTON; QUANTUM CHROMODYNAMICS; QUARKS; STRONG NUCLEAR INTERACTIONS.

$$\begin{aligned} \text{Quark} + \text{quark} &\rightarrow (\text{quark} + \text{gluon}) + \text{quark} \\ &\rightarrow \text{quark} + (\text{gluon} + \text{quark}) \\ &\rightarrow \text{quark} + \text{quark} \qquad (8) \end{aligned}$$

Gauge interactions. Three of the four fundamental interactions (electromagnetic, weak nuclear, and strong nuclear) appear to be mediated by intermediate quanta (photons γ; W^+, Z^0, and W^-; and gluons G, respectively), each carrying spin of magnitude $\hbar$. This is characteristic of the gauge interactions, whose general theory was given by H. Weyl, C. N. Yang, R. Mills, and R. Shaw. This class of interactions is further characterized by the fact that the force between any two particles (produced by the mediation of an intermediate gauge particle) is universal in the sense that its strength is (essentially) proportional to the product of the intrinsic charges (electric, or weak-nuclear, or strong-color) carried by the two interacting particles concerned.

The fourth interaction (the gravitational) can also be considered as a gauge interaction, with the intrinsic charge in this case being the mass; the gravitational force between any two particles is proportional to the product of their masses. The only difference between gravitation and the other three interactions is that the gravitational gauge quantum (the graviton) carries spin $2\hbar$ rather than $\hbar$. As discussed below, it is an open question whether all fundamental interactions are gauge interactions. *See* GAUGE THEORY.

Unification of Interactions

Ever since the discovery and clear classification of these four interactions, particle physicists have attempted to unify these interactions as aspects of one basic interaction between all matter. The work of M. Faraday and J. C. Maxwell in the nineteenth century, which united the distinct forces of electricity and magnetism as aspects of a single interaction (the gauge interaction of electromagnetism), has served as a model for such unification ideas.

Gravitation and electromagnetism. The first attempt in this direction was made by Einstein who, having succeeded in understanding gravitation as a manifestation of the curvature of space-time, tried to comprehend electromagnetism as another geometrical manifestation of the properties of space-time, thus achieving a unification between these forces. In this attempt, to which he devoted all his later years, he is considered to have failed. *See* RELATIVITY.

Electroweak interaction. A unification of weak and electromagnetic interactions, employing the gauge ideas discussed above, was suggested by S. Glashow and by A. Salam and J. C. Ward in 1959. This followed a parallel between these two interactions, pointed out by J. S. Schwinger in 1957. Assuming that (the then known) weak interactions (4) were mediated by exchanges of (the then hypothetical) W^+ and W^- particles, it could be shown from the empirical properties of weak interaction phenomena, that if the W's existed, they must carry an intrinsic spin of magnitude $\hbar$, just as does the photon, the gauge quantum of electromagnetism. If a bold unifying assumption was made that this magnitude of spins $\hbar$ for W^+, W^-, and the photon γ connotes a gauge character for a unified electroweak interaction, and that the intrinsic coupling strength of weak interactions is universally the same as that for electromagnetism (that is, $\alpha = 1/137$), then it could be shown that the masses of the W^+ and W^- particles must be in excess of the quantity given in Eq. (9).

$$\sqrt{\pi\alpha\hbar^3/(\sqrt{2}G_F c)} = 37.4\ \text{GeV}/c^2 \qquad (9)$$

Following this initial attempt, Glashow (and independently Salam and Ward) noted that such a unification hypothesis is incomplete, inasmuch as electromagnetism is a left-right symmetry-preserving interaction, in contrast to the weak interaction, which violates this symmetry. A gauge unification of such disparate interactions could be effected only if, additionally, new weak interactions represented by reactions (5) are also postulated to exist. Equivalently, there must exist a new electrically neutral intermediate weak-quantum Z^0 besides the (hypothetical) W^+ and W^-.

Spontaneous breaking and renormalization. There were two major problems with this unified electroweak gauge theory considered as a fundamental theory. Yang and Mills had shown that masslessness of gauge quanta is the hallmark of unbroken gauge theories. The origin of the masses of the weak interaction quanta W^+, W^-, and Z^0 (or equivalently the short-range of weak interactions), as contrasted with the masslessness of the photon (or equivalently the long-range character of electromagnetism), therefore required explanation. The second problem concerned the possibility of reliably calculating higher-order quantum effects with the new unified electroweak theory, on the lines of similar calculations for the "renormalized" theory of electromagnetism elaborated by S. Tomonaga, Schwinger, Feynman, and F. J. Dyson around 1949. The first problem was solved by S. Weinberg and Salam and the second by G. t'Hooft and by B. W. Lee and J. Zinn-Justin. *See* RENORMALIZATION.

Weinberg and Salam considered the possibility of the electroweak interaction being a "spontaneously broken" gauge theory. By introducing an additional self-interacting Higgs-Englert-Brout-Kibble particle into the theory, they were able to show that the W^+, W^-, and Z^0 would acquire well-defined masses through the so-called Higgs mechanism, these masses being given by Eqs. (10), where 37.4 GeV/c^2

$$m_W = \frac{37.4\ \mathrm{GeV}/c^2}{\sin\theta_w} \qquad m_Z = \frac{37.4\ \mathrm{GeV}/c^2}{\sin\theta_w \cos\theta_w} \tag{10}$$

is the combination of constants given by Eq. (9). Here θ_w is a weak mixing parameter for electromagnetism and weak interactions. The constant $\sin^2\theta_w$ can be determined from experiments which give the ratios of cross sections of Z^0-mediated reactions (6) to the W^+ and W^--mediated reactions (4). The best available value, calculated from all low-energy experiments, is given by Eq. (11).

$$\sin^2\theta_w \approx 0.230 \tag{11}$$

See SYMMETRY BREAKING.

The predicted theoretical mass values of the W and Z particles deduced by substituting Eq. (11) into Eqs. (10) are in good accord with the experimental values found by the CERN 1983 experiments. The existence of the W and Z particles and this accord with regard to mass values give support to the basic correctness of the electroweak unification ideas, as well as to the gauge character of the electroweak interaction.

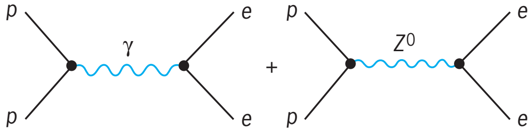


Fig. 2. Feynman diagrams corresponding to the Stanford (1978) experiment, which showed the interference effects between mediating photons and Z^0 exchanges in electron-proton scattering.

Prior to this direct evidence, indirect evidence for the existence of the characteristic reactions (6), predicted by the electroweak theory, had existed since 1973. The most crucial experiment in this respect, carried out at Stanford during 1978, exhibited interference effects between the photon (γ) and the Z^0 particle in the scattering of polarized electrons from protons (**Fig. 2**). These effects were established through observing the characteristic weak left-right symmetry violation in the reaction $e^- + d \rightarrow e^- + d$. The findings of this experiment provided indirect but quantitative confirmation of the predictions of the electroweak theory.

Higgs particle. The Weinberg-Salam electroweak theory contains an additional neutral particle (the Higgs) but does not predict its mass. A search for this particle will be seriously undertaken when the electron-positron accelerators LEP (Large Electron-Positron storage ring) at CERN and SLC (Stanford Linear Collider) come into commission. Meanwhile, there has been theoretical speculation on whether the Higgs particle is a composite object held together by a new, fundamental type of very strong interaction, the so-called technicolor interaction. This suggestion has the possible merit of eliminating the need for introducing a fundamental nongauge (self-) interaction among Higgs particles. *See* ELECTROWEAK INTERACTION; HIGGS BOSON; PARTICLE ACCELERATOR; STANDARD MODEL.

Electronuclear interaction. The gauge unification of weak and electromagnetic interactions, which started with the observation that the relevant mediating quanta (W^+, W^-, Z^0, and γ) possess intrinsic spin $\hbar$, can be carried further to include strong nuclear interactions as well, if these strong interactions are also mediated through quanta (gluons) carrying spin $\hbar$. The resulting theory, which appears to explain all known low-energy phenomena, is called the standard model. (It is a model based on three similarly constituted generations of quarks and leptons plus the mediating quanta W^+, W^-, Z^0, photons, and gluons plus the Higgs particle.) A complete gauge unification of all three forces (electromagnetic, weak-nuclear, and strong-nuclear) into a single electronuclear interaction seems plausible. Such a (so-called grand) unification necessarily means that the distinction between quarks on the one hand and neutrinos, electrons, and muons (leptons) on the other, must disappear at sufficiently high energies, with all interactions (weak, electromagnetic, and strong) clearly manifesting themselves then as facets of one universal gauge force with a primitive universal strength equal to $\alpha/\sin^2\theta_w$. The fact that at low energies presently available, these interactions exhibit vastly different effective strengths is ascribed to differing renormalizations due to successive spontaneous symmetry breakings. A startling consequence of the eventual universality and the disappearance of distinction between quarks and leptons is the possibility, first discussed by J. C. Pati and Salam within their electronuclear model, of protons transforming into leptons and pions. Contrary to the older view, protons would therefore decay into leptons and pions and not live forever. A somewhat different model elaborated a year later by H. Georgi and Glashow (called the grand unifying theory), predicts a lifetime of the order of 10^{29} years for the proton P, with decay principally through the mode $P \rightarrow e^+ + \pi^0$, where π^0 is the neutral pion and e^+ is the positron. Experiments carried out during 1983 to search for this mode of proton decay gave negative evidence for protons decaying at this rate, although other types of decay modes may have been observed in other experiments. The discovery of proton instability (with decays into leptons or antileptons) would be an epic discovery and a direct confirmation of the electronuclear (grand) unification. *See* GRAND UNIFICATION THEORIES; PROTON.

Consequences of symmetry breaking. Spontaneous symmetry breaking of gauge interactions has the characteristic that symmetry breaking is a phase phenomenon and disappears in a high-temperature environment. This implies that, at temperatures T in excess of 10^{15}K (T greater than $m_z c^2/k$, where k is the Boltzmann constant), that is, up until 10^{-12} s after the outset of the big bang, there was no spontaneous breaking of the symmetry of electroweak interactions, and the W and the Z particles were massless, like the photons and the gluons. The onset of such phase transitions plays a crucial role in modern cosmological theories of the early universe, resolving some old dilemmas. For example, proton decay, and left-right and particle-antiparticle symmetry violations, provide a natural explanation for the fact that the present universe contains a preponderance of protons and neutrons rather than of their antiparticles. However, the existence of such phase transitions also poses some new dilemmas, such as the prediction of the existence of heavy magnetic monopoles (in the early universe), with abundances surviving into the present epoch, for which there is no experimental evidence. To remedy this, it is necessary to postulate an inflationary epoch having occurred in the universe's history somewhere about 10^{-33} s after the onset of the big bang. *See* ANTIMATTER; BIG BANG THEORY; COSMOLOGY; INFLATIONARY UNIVERSE COSMOLOGY; MAGNETIC MONOPOLES; PHASE TRANSITIONS; UNIVERSE.

Prospects for including gravity. Research in unification theories of fundamental interactions is now concerned with uniting the gauge theories of gravity and of the electronuclear interactions. One promising approach is the extension of space-time to more than four dimensions, following ideas developed by T. Kaluza and O. Klein in the 1920s. Remarkably, the formal expression for Einstein's gravitational interaction in a space-time of dimensions higher than four, is equivalent to the standard Einstein theory of spin-$2\hbar$ gravitons in four dimensions plus a Yang-Mills theory of spin-$\hbar$ particles (that is, a theory describing the electronuclear type of gauge interactions) when the extra dimensions are contracted down to less than 10^{-35} m. No realistic model of such a compactified unified theory has emerged, though Einstein-like supersymmetric theories in 10-space and 1-time (a total of 11 dimensions) are the favored candidates. (Supersymmetry is the principle which treats gauge and Higgs particles on a par with quarks and leptons.) *See* SUPERGRAVITY; SUPERSYMMETRY.

The most promising approach appears to be that of superstring theories. Such theories appear to describe the only possible theory of gravity which is finite and suffers from no ultraviolet infinities. A closed string is a (one-dimensional) loop which may exist in a d-dimensional space-time (where d must equal 10 to completely eliminate all ultraviolet infinities). The quantum oscillations of the string correspond to particles of higher spins and higher masses, which may be strung on a linear trajectory in a spin-versus-mass^2 (Regge) plot. Among these are the zero-mass gravitons and the gauge mesons. The theory has a unique built-in gauge symmetry.

So far, it has not been possible to go down from $d = 10$ dimensions to $d = 4$ conventional space-time dimensions and to produce the emergence of the standard model, although there is hope that this may be accomplished. If these string ideas are successful, they may help lead to one single theory which unites all known low-energy phenomena. *See* SUPERSTRING THEORY.

Abdus Salam

Bibliography. R. A. Carrigan, Jr., and W. P. Turner, *Fundamental Particles and Forces*, 1989; S. Glashow, A. Salam, and S. Weinberg, Nobel Lectures in Physics, *Rev. Mod. Phys.*, 52:515–543, 1980; G. Kane, *Modern Elementary Particle Physics: The Fundamental Particles and Forces*, 1993; C. Quigg, Elementary particles and forces, *Sci. Amer.*, 252(4): 84–95, April 1985; C. Rubbia, Experimental observation of the intermediate vector bosons W^+, W^- and Z^0, Nobel Lecture, December 8, 1984, *Rev. Mod. Phys.*, vol. 57, no. 3, part 1, pp. 699–722, 1985.

Fungal biotechnology

All aspects of cultivating fungi together with products and processes derived from such cultures. Fungi, including yeasts, are an extremely diverse group of heterotrophic organisms with an enormous impact, both beneficial and detrimental, on the environment, on animal and plant life, and on the global economy and human social development. Even though they are unable to synthesize their own food, they exhibit a wide range of biosynthetic and biodegradative activities. Since fungi can bring about chemical change in almost any natural or synthetic organic molecule, many species have been selected and propagated in pure culture specifically for applications in biotechnology and industry, for example in food and beverage production.

While the fermentation industry remains the largest and economically most important user of fungal cultures, fungi are also utilized by the pharmaceutical, cosmetic, chemical, agricultural, food, enzyme, wood product, and waste treatment industries. In the United States, two prominent examples of fungal products are citric acid, with an annual production of 350,000 tons (160,000 kg), and beta-lactam antibiotics. Yeasts are the most commercially exploited microorganisms. They have been used extensively in baking, brewing, winemaking, and distilling, and in making various metabolic products. Several million tons of fresh yeast are produced each year, mostly for the baking industry. *See* FERMENTATION; FOOD FERMENTATION.

Fungi stand at the forefront of biotechnology. The technology for large-scale fermentation and downstream processing is well understood. Newer techniques, such as protoplast fusion and genetic engineering, along with immobilized molecules and powerful fermentation processes, will lead to further utilization of the biosynthetic potential of fungi. Moreover, any fungus that can be grown in culture is a prospective host for the production of economically important, biologically active proteins native to higher forms.

Alcohols. The production of alcohols by fungi through the fermentation of sugar is the oldest industrial fermentation process. Of the hundreds of known metabolic products of fungi, ethanol is perhaps the most widely used. Ethanol is an ideal fuel and the basis for countless chemical transformations. It serves as a solvent for different materials and as the raw material from which other compounds are synthesized. As a solvent, it is employed in the manufacture of such diverse materials as dyes, drugs, soaps, plastics, resins, and polishes; and as a raw material, in the synthesis of compounds such as ether, esters, acetic acid, hypnotics, and synthetic rubber.

Ethanol (ethyl alcohol) from the yeast *Saccharomyces cerevisiae* is the largest fungal fermentation product in both volume and value. It is used primarily in the production of alcoholic beverages such as beer, wine, distilled spirits, cider, sake, and liqueurs. Fuel ethanol produced by yeast fermentation of sugarcane juice and molasses is used as an octane enhancer and replacer in fuel (gasoline) for internal combustion engines in the majority of automobiles in Brazil. Other industrially useful alcohols produced by yeast fermentation are higher alcohols, and polyhydric alcohols such as glycerol, xylitol, sorbitol, arabinitol, erythritol, and mannitol. Glycerol is widely used in the production of synthetic resins, pharmaceuticals, cosmetics, and toothpastes. However, the current production of glycerol by fermenation is insignificant compared with chemically produced glycerol. Xylitol from *Candida* species finds increasing uses in the food industry as a sweetener and as a sugar substitute for insulin-dependent diabetics. Arabinitol from *Candida entomaea* is used as a low-calorie bulking agent, in biosurfactants, as an edible coating, and as a carrier for pharmaceutical agents. *See* ALCOHOL; ALCOHOL FUEL; ETHYL ALCOHOL; FOOD MANUFACTURING.

Amino acids. Amino acids are the building blocks of proteins. While bacteria are more likely to be considered as amino acid–producing organisms because of their rapid growth rate, fungi represent the simplest eukaryotic model system that possesses the full capacity of amino acid synthesis. Fungi are actually superior to bacteria in their ability to grow on a variety of polysaccharides, wood wastes, or other cheap materials. Their potential to produce amino acids may become even more important if such carbon sources are used.

Amino acids from fungi are used as food supplements and precursors for products. So far, however, only tryptophan, lysine, and methionine have been produced commercially. Phenylalanine is used in the synthesis of the artificial sweetener aspartame. L-Lysine is known to be produced by *Apiotrichum humicola*, *Aspergillus ustus*, *Candida lipolytica*, *Candida pelliculosa*, *Cryptococccus humicolus*, *Cryptococccus laurentii*, *Kloeckera africana*, *Sphacelotheca sorgii*, *Trichosporon beigelii*, and *Ustilago maydis*; L-methionine by *Aspergillus oryzae*, *Candida lipolytica*, *Candida sake*, and *Hanseniaspora valbyensis*; L-phenylalanine by *Rhodotorula glutinis*; and L-tryptophan by *Candida sake*, *Candida zeylanoides*, *Claviceps fusiformis*, *Claviceps purpurea*, *Hansenula anomala*, *Pichia anomala*, and *Pichia pastoris*. *See* AMINO ACIDS.

Enzymes. Fungal sources currently provide the largest assortment of bulk commercial enzymes with the widest variety of applications. In some processes the crude enzymes secreted by cells to act upon their substrates in an external environment can be exploited without the preparation of refined enzyme extracts.

Fungal enzymes range from bulk technical, detergent, and food grade (typically proteases and carbohydrate-hydrolyzing enzymes available at low cost and in less than two dozen general types) to the specialty enzymes used for research or analytical needs. The latter are available in many varieties and forms and are high in purity as well as cost.

The food industry alone consumes more than half the world's output of industrial enzymes. The majority of food enzymes are hydrolases, such as proteases from *Aspergillus oryzae* for bread-making and cheese-making, pectinases from *Penicillium* and *Aspergillus* for fruit processing, and amylases from *Aspergillus* for the brewing industry. *Trichoderma viride* and *Aspergillus niger* are commercial sources of cellulases for treating various feedstuffs and food modifications. *See* CHEESE; ENZYME; MALT BEVERAGE.

For the manufacture of paper pulp, lignin can be removed from wood without destroying the cellulose and hemicellulose by the lignin-degrading enzymes of certain white-rot fungi, such as *Phanerochaete chrysosporium*. (Brown-rot fungi utilize only cellulose and cannot degrade lignin.) Biological delignification has the advantages of low capital cost, low energy input, greater specificity, and high yield without generating polluting by-products. In addition, nonspecific lignases of *Pichia chrysosporium* can degrade a variety of phenol-containing compounds, such as pesticides, and wastewater discharges containing trinitrotoluene. The wastewater from industries that use molasses as a raw material contains a dark brown pigment that is not decolorized by usual biological treatments. The phenol-oxidizing enzymes of *Aspergillus oryzae*, *Pleurotus*, or *Auricularia* can decolorize such fermentation wastes. Decolorization of sugarcane molasses is also accomplished in the food industry by fungi. Among the yeasts, *Lipomyces* can degrade the herbicide paraquat, *Rhodotorula* can convert benzaldehyde to less toxic benzyl alcohol, and *Candida* can degrade formaldehyde. Species of *Pleurotus* are able to break down pentachlorophenol (PCP) and DDT. The filamentous fungi *Aspergillus niger* and *Chaetomium cupreum* degrade the tannins in tannery effluents. White-rot fungi, such as the mushroom *Coriolus versicolor*, are effective in solubilizing low-grade coal, which has a structure similar to lignin. For retting flax fibers (loosening the fibers from the woody tissue), the polysaccharide-degrading enzymes of the mushroom *Pleurotus* can be used instead of boiling with sodium hydroxide for removal of noncellulosic materials. Enzyme-treated fibers produce high-quality yarns. A lipase from *Thermomyces lanuginosus* is used to remove fatty stains such as those made by olive oil and lard from clothes. A commercial fungal cellulase is used to remove the microfibers during washing without degrading the cotton fibers, making the cotton fibers regain the smooth appearance characteristic of new fabrics and making the fabric stay soft and bright in color longer than when washed without the cellulase. *See* PAPER.

Organic acids. Fungi produce over 40 different organic acids. The compounds may accumulate as the result of carbohydrate metabolism, they may be terminal products of glycolysis, or they may result from the incomplete oxidation of sugars. However, only citric acid, itaconic acid, and gluconic acid are produced in significant quantities by large-scale commercial processes. Citric acid is widely used in the food industry for fruit drinks, confectionery, jams, and preserved fruit, and in the cosmetic industry for astringent lotions. The ability of citric acid to sequester metal ions in acidic media makes it useful in the formulation of shampoos and electrolyte solutions, in the tanning of animal skins, and even for the clearing of pipes in the oil industry. It is produced industrially almost exclusively by *Aspergillus niger* and *Aspergillus wentii*. Gluconic acid from *Aspergillus niger* is used in medicine as a diet supplement. Itaconic acid from *Aspergillus terreus* is used in the production of fibers, paint adhesives, and thickening agents. *See* CITRIC ACID.

Due to the acids they produce, fungi are successful bioleaching agents. For example, *Aspergillus niger*, *Scopulariopsis brevicaulis*, and *Penicillium expansum* are capable of extracting aluminum from aluminosilicates. *Penicillium* can leach out nickel and insoluble phosphates from ore. Yeasts have been used to leach zinc, copper, and lead from filter dusts from copper-process converters.

Polysaccharides. Fungi produce a number of nonprotein polymers. Almost all are polysaccharides composed of simple sugars. Industrial applications for the polymers are numerous and varied: stabilizers for such food products as syrups and ice cream; pigment suspension in paint of high viscosity; dispersion of pigments for fabric printing and dying; additives to drilling mud in the oil exploration industry; extraction of crude oil from oil-bearing shales; improvement of the viscosity of cosmetic lotions and dentifrices. They are also used as insoluble matrices for immobilizing enzymes and cells. An advantage of using fungi for the production of polymers is their ability to grow at low pH, limiting most bacterial contaminations. Thus, growth conditions for large-scale production do not have to be rigorously aseptic. *See* POLYSACCHARIDE.

Pullulan produced by *Aureobasidium pullulans* is a neutral water-soluble polysaccharide. It is colorless, odorless, tasteless, nontoxic, low-calorie, completely biodegradable, and highly soluble in water; it adheres to surfaces such as paper, wood, glass, or metal. One application is as an adhesive in the manufacture of wood-based building products. It can also be formed into films, fibers, or shaped articles.

Some of the more recently isolated substances are metabolic regulators, antitumor drugs, and immunomodulators that enhance host resistance against cancer, immunodeficiency disease, or generalized immunosuppression. Four commercially available polysaccharide preparations are currently used in treating various cancers: lentinan from *Lentinus edodes*, schizophyllan (SPG) from *Schizophyllum commune*, the polysaccharide-polypeptide PSK (krestin) or PSP from *Coriolus versicolor*, and scleroglucan from *Sclerotium rolfsii* and *Sclerotium glucanicum*. The yeasts *Hansenula*, *Pichia*, and *Pachysolen* produce extracellular phosphorylated mannans useful as thickening, stabilizing, and dispersing agents. Levans from yeasts and *Aspergillus sydowii* and *Aspergillus versicolor* can substitute for gum arabic. Chitosan extracted from *Mucor rouxii* and *Absidia coerulea* is used as a flocculating agent for clarifying sewage, in oil recovery, as a paper-sizing agent, and as a chelating agent for metal ions.

Food proteins. Fungi can grow on almost anything that contains carbohydrates, including vegetable and distillery wastes and wood-processing effluents. Because of their rapid growth rate and high protein content, fungi are an ideal source of protein for animals and humans. Modern production of single-cell protein from yeasts, such as *Candida utilis*, is making fungal protein competitive with other more traditional animal feeds. Mycoprotein, a product from filamentous fungi, is used as a meat substitute.

At present, mushroom culture represents the only major process in biotechnology that can successfully convert industrial waste, wood waste, and urban refuse into a product that is directly edible by humans. The most common mushroom grown for commercial purposes in the United States is the white button species, *Agaricus bisporus*. In Asian countries, shiitake (*Lentinula edodes*), the padi straw mushroom (*Volvariella volvacea*), and the oyster fungus (*Pleurotus ostreatus*) are popular and profitable crops. *See* MUSHROOM.

Flavor, aroma, and coloring compounds. While fungi have long been utilized in the food industry, their potential for the production of natural flavor and color compounds has been recognized only recently. In comparison to expensive extraction from plants or chemical synthesis, microbial production of aromatic compounds offers definite advantages, particularly for consumers wishing to avoid chemical additives in food and other household products. Certain filamentous fungi have been used to produce desirable flavors in cheeses, such as *Penicillium roquefortii* and *Penicillium camembertii*, and in the manufacture of fermented foods, particularly in Asian countries. Examples of popular fermented foods are soy sauce and miso, which utilize the *Aspergillus flavus-oryzae* group of fungi, and tempeh from the action of *Rhizopus*. A variety of yeasts and filamentous fungi, such as *Penicillium notatum*, can produce sweet, fruity, coconutlike, buttery, or nutlike flavors. In addition, a wide assortment of compounds with unique and potent flavors have been isolated as naturally occurring metabolites of edible and wood-decaying mushrooms. The filamentous fungus *Monascus purpureus* has been used in the production of red color in rice wine. Both yeasts and filamentous fungi can produce orange-colored β-carotene, which is a popular, nontoxic additive to such foods as butter, ice cream, orange juice, and candies. The carotenoid astaxanthin from the red yeast *Phaffia rhodozyma* is used as a pigment in salmonid feeds in fish farms, and is efficiently absorbed in the flesh. *See* SPICE AND FLAVORING.

Vitamins and plant growth factors. Fungi can produce a variety of vitamins, including biotin, folic acid, pantothenic acid, pyridoxine, riboflavin, thiamine, and cyanocobalamin, but only the production of riboflavin has been commercialized. Riboflavin from *Acremonium diospyri*, *Ashbya gossypii*, *Candida guilliermondii*, *Eremothecium ashbyi*, *Nematospora gossypii*, and *Pachysolen tannophilus* competes well with chemical synthesis. In addition, *Aspergillus giganteus*, *Blakeslea trispora*, *Mortierella ramanniana*, and *Phycomyces blakesleanus* produce β-carotene, a precursor of vitamin A, and *Candida norvegensis* produces ascorbic acid, vitamin C. *See* VITAMIN.

Plant growth regulators are chemical compounds with the capability of coordinating the growth and morphogenesis of plants. Gibberellin is the only one produced commercially by means of fungal fermentation; *Fusarium moniliforme* and *Gibberella fujikuroi* are used. Under the right conditions, some strains of *Gibberella fujikuroi* yield only one gibberellin (gibberellic acid, or gibberellin A3), which in malting is sometimes applied to the germinating barley to increase the yield of amylases. Gibberellic acid has received much experimental attention in horticulture for the induction of flowering of plants independently of photoperiodism. Its practical use so far has been limited to a few applications, such as obtaining seedless grapes. *See* GIBBERELLIN.

Antibiotics and drugs. Antibiotics are among the most important compounds produced by fungi. The timely discovery and successful use of penicillin derived from *Penicillium chrysogenum* during World War II provided an early model to guide production of new antibiotics. Penicillin is not one chemical compound but a mixture of several, each with varying therapeutic values. Cephalosporin C, a structurally related antibiotic produced by *Acremonium*, is the parent compound for all of the cephalosporins. Fusidic acid, which was first isolated from *Fusidium coccinellum*, is an antibiotic effective against gram-positive infections, particularly penicillin-resistant staphylococci. Griseofulvin is an antifungal antibiotic produced by *Penicillium griseofulvin*. It is the only effective antibiotic available for the systemic treatment of fungal infections of skin, hair, and nails. *See* ANTIBIOTIC.

The cyclosporins from *Tolypocladium inflatum* and *Trichoderma polysporum*, first identified as antifungal antibiotics, also exhibit immunosuppressive activity. Their use has greatly reduced morbidity and increased survival rates in human transplant patients. Gliotoxin from *Aspergillus fumigatus* has antibiotic and immunosuppressive properties.

More than 40 alkaloids produced by *Claviceps purpurea* have been isolated from ergot, and hundreds of derivatives have been chemically prepared. Alkaloids have found use in the treatment of migraine headaches, in the control of uterine motor activity, and in the therapy of Parkinson's disease. Asperlicin from *Aspergillus allicaceus*, an antagonist of cholecystokinin, is used in the treatment of disorders of the gastrointestinal system that involve cholecystokinin. Lovastatin from *Aspergillus terreus* and *Monascus ruber* is a very potent inhibitor of cholesterol biosynthesis and is used for lowering cholesterol levels. Mevastatin (compactin) from *Hypomyces chrysospermus*, *Penicillium brevicompactum*, *Penicillium citrinum*, *Trichoderma longibrachiatum*, and *Trichoderma pseudokoningii* and paravastatin from *Mucor hiemalis* also lower plasma cholesterol levels. *See* ALKALOID.

Recombinant proteins. Because they are capable of secreting large quantities of certain proteins in liquid culture, fungi have proven to be useful as cloning hosts for the production of recombinant proteins of fungal and human origin. Fungi commonly used as transformation agents are *Saccharomyces cerevisiae*, *Aspergillus awamori*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, and *Pichia pastoris*. Technologies have been developed to scale up the production of new or novel products. *Aspergillus nidulans* has been designed to produce many human therapeutic proteins, including growth factors and protein hormones. The yeasts *Saccharomyces cerevisiae* and *Pichia pastoris* have been utilized for the expression of human interferon and serum albumin. Yeast chromosomes are also being employed in the mapping of the human genome. *See* GENETIC ENGINEERING.

Shung-Chang Jong

Mycopesticides. Manufactured products which contain living fungi and are used to control pests are called mycopesticides. They are utilized to control weeds, harmful insects, nematodes (roundworms), or even other fungi. Although formerly confined to experimental settings, mycopesticides are increasingly available as commercial products, especially for the agricultural market. They may impact pest populations through direct parasitism, secretion of antibiotics, competition for nutrients, or a combination of these effects, and may be used alone or in combination with chemical pesticides.

The fungi used in formulating mycopesticides are usually selected from natural populations for their ability to survive environmental conditions, control the target pest, and avoid damage to nontarget organisms. However, experimental genetic manipulation has been employed to enhance one or more desirable properties of these fungi, such as rendering the incorporated fungus resistant to a particular fungicide.

Any commercially available mycopesticide must be formulated to withstand conditions of storage, transport, and application; must be effective against the target pest; and must be thoroughly tested for safety for humans and other nontarget organisms. Mycopesticides containing genetically modified organisms are subject to additional regulatory controls. In the United States, the principal regulatory authority for pesticides is the Environmental Protection Agency. If genetically modified microorganisms are released in the field, the Animal and Plant Health Inspection Service (U.S. Department of Agriculture) has additional oversight.

Examples of fungi that have been utilized in mycopesticide production include *Colletotrichum gloeosporioides* f.sp. *aeschonomene* (against northern jointvetch), *Puccinia chondrillina* (against skeletonweed), and *Entyloma ageratinae* (against mistflower); *Metarhizium anisopliae* (against cockroaches and termites), *Beauveria bassiana* (against grasshoppers, locusts, and beetles), *Verticillium lecanii* (against whiteflies and thrips), and *Paecilomyces fumosoroseus* (against whiteflies); *Arthrobotrys superba* (against plant-parasitic nematodes); *Gliocladium virens* (against the fungi *Rhizoctonia* and *Pythium*), and *Trichoderma harzianum* (against wood rot fungi). *See* FUNGI; FUNGISTAT AND FUNGICIDE; PESTICIDE. Frank M. Dugan

Bibliography. A. Altman (ed.), *Agricultural Biotechnology*, Marcel Dekker, New York, 1998; D. K. Arora, R. P. Enlander, and K. G. Mukerji, *Handbook of Applied Mycology*, vol. 4: *Fungal Biotechnology*, 1992; M. A. Levin and E. Israeli (eds.), *Engineered Organisms in Environmental Settings: Biotechnological and Agricultural Applications*, CRC, New York, 1996; R. G. Van Driesche and T. S. Bellows, Jr., *Biological Control*, Chapman & Hall, New York, 1996; M. Wainwright, *An Introduction to Fungal Biotechnology*, 1992; G. M. Walker, *Yeast Physiology and Biotechnology*, 1998; D. T. Wicklow and B. E. Söderström (eds.), *The Mycota*, vol. 4: *Environmental and Microbial Relationships*, Springer-Verlag, Berlin, 1997.

Fungal ecology

The subdiscipline in mycology and ecology that examines community composition and structure; responses, activities, and interactions of single species; and the functions of fungi in ecosystems. The fungi, as they have been historically classified, comprise organisms from two distinct evolutionary lineages: the true fungi, or Eumycota; and the Oomycota, which are allied with chrysophyte algae in the separate kingdom Chromista. The oomycetes resemble the true fungi in having a filamentous growth form and occurring as plant parasites or free-living saprobes (organisms living on decaying organic matter) in soil or decaying vegetation. The slime molds, or Mycetozoa, are another distinct group in the kingdom Protista. These organisms display an extraordinary diversity of ecological interactions and life history strategies, but are alike in being efficient heterotrophs. In all ecosystems, producers (generally plants and some bacteria) synthesize cell components by using solar energy to chemically reduce carbon dioxide and assemble carbohydrates into the complex polymers of cell walls. Heterotrophs consume and disassemble, obtaining energy from the catabolism of photosynthetically formed carbohydrates and other molecules. Fungi and bacteria are the primary decomposers, facilitating the flow of energy and the cycling of materials through ecosystems. Both classes of organisms occur wherever there is organic matter. With an estimated 1.5 million species worldwide, of which only about 5% have actually been named, the fungi rank just below the insects in biodiversity. *See* ECOLOGICAL ENERGETICS; FUNGI.

Habitats. Fungi occur in many different habitat types—on plant surfaces; inside plant tissues; in decaying plant foliage, bark and wood; and soil—generally changing in abundance and species composition through successional stages of decomposition. Fungi are also found in marine and aquatic habitats; in association with other fungi, lichens, bacteria, and algae; and in the digestive tracts and waste of animals. Some fungi grow in extreme environments: rock can harbor free-living endolithic fungi or the fungal mutualists of lichens; thermotolerant and thermophilic fungi can grow at temperatures above 45°C (113°F); psychrophilic fungi can grow at temperatures to below −3°C (27°F). Xerotolerant fungi are able to grow in extremely dry habitats, and osmotolerant fungi grow on subtrates with high solute concentrations. Most fungi are strict aerobes, but species of the chytrid *Neocallimastix*, which inhabit the rumen of herbivorous mammals, are obligate anaerobes. Several aquatic fungi are facultative anaerobes.

Although fungi are ubiquitous in terrestrial habitats, their abundance and diversity in these various domains is highly variable. Many fungi occur as free-living saprobes, but fungi are particularly successful as mutualistic, commensal, or antagonistic symbionts with other organisms. Fungi are ecologically and economically important as the principal pathogens of vascular plants. However, they are also beneficial mutualistic mycorrhizal associates of roots, which aid in nutrient acquisition. Fungi are the mutualistic partners with algae or cyanobacteria in lichens, and endobiotic fungi in certain grasses produce toxins that inhibit insect herbivores. Fungal associations with nematodes, protozoa, mites, and insects also range from predation to mutualism. Fungi are cultivated and consumed as food by leaf-cutting ants, termites in the family Macrotermitinae, and wood-boring ambrosia beetles. Several groups of fungi have evolved a variety of specialized means for capturing or parasitizing nematodes and protozoa. The oomycete *Haptoglossa* pierces the body of rotifers with a harpoonlike organelle. Predacious fungi capture rotifers and nematodes by adhesive knobs and nooselike constricting hyphal rings. Commensal trichomycetes inhabit the digestive tracts of arthropods. Fungal parasitism is widespread among several groups of fungi specializing on various insects, such as *Cordyceps* on Lepidoptera, *Entomophthora muscae* on houseflies, and *Septobasidium* on scale insects. Laboulbeniales, an order with over 1800 species, are specialized insect parasites that occur on several insect families, especially beetles. Fungi also are the principal or exclusive diet of species, genera, and sometimes families of protozoa, nematodes, mites, collembola, and insects. Truffles and mushrooms are an important component of the diets of several small mammals. *See* POPULATION ECOLOGY.

Biomass. Fungal hyphae are pervasive in soil and decomposing plant litter. The lengths of hyphae in deciduous forest materials have been estimated at about 1300 ft (400 m) per gram of dry soil and up to 20 times that length in the humus or litter above. In coniferous forest systems, the comparable figures are 1–1.6 mi (1500–2500 m) per gram of dry soil [37 mi (60,000 m) in mycorrhizal mats] and about 41.5 mi (67,000 m) per gram of dry organic horizon. Fungal biomass in tropical soils may be twice that in temperate zone soils. Fungal individuals, as a contiguous, genetically distinct hyphal network (thallus), are among the largest and oldest organisms. A single fungus of the tree root colonizer *Armilliaria gallica* was estimated to span 2083 ft (635 m), to weigh over 22,000 lb (10,000 kg) and to be over 1500 years old. In contrast, genets (individuals that are genetically distinct) of parasitic or saprobic fungi may consist of only a few cells.

Significant features. Fungi possess unique features that affect their capacity to adapt and to function in ecosystems:

1. Fungi are composed of a vegetative body (hyphae or single cells) capable of rapid growth. Hyphae are linear strands composed of tubular cells that are in direct contact with the substrate. The cells secrete extracellular enzymes that degrade complex polymers, such as cellulose, into low-molecular-weight units that are then absorbed and catabolized. Many fungi also produce secondary metabolites such as mycotoxins and plant growth regulators that affect the outcomes of their interactions with other organisms.

2. Filamentous fungi are able to mechanically penetrate and permeate the substrate.

3. Fungi have an enormous capacity for metabolic variety. Fungal enzymes are able to decompose highly complex organic substances such as lignin, and to synthesize structurally diverse, biologically active secondary metabolites. Saprotrophic fungi are very versatile; some are able to grow on tree resins and even in jet fuel.

4. Structural and physiological features of fungi facilitate absorption and accumulation of mineral nutrients such as phosphorus, nitrogen, and potassium, as well as toxic elements such as arsenic, selenium, and mercury. The capacity of fungi to absorb, accumulate, and translocate is especially significant ecologically where hyphal networks permeate soil and function as a link between microhabitats. For example, mycorrhizal fungi serve as a link between soil and their plant symbionts, an interconnected network allowing movement of carbon and minerals.

5. Fungi have the capacity for indeterminate growth, longevity, resilience, and asexual reproduction. The vegetative cells of Eumycota are often multinucleate, containing dissimilar haploid nuclei. This combination of features gives the fungi an unparalleled capacity for adaptation to varying physiological and ecological circumstances and ensures a high level of genetic diversity.

6. Many species of fungi have a capacity to shift their mode of nutrition. The principal modes of nutrition in fungi are saprotrophy (the utilization of dead organic matter) and biotrophy, which is characteristic of parasitic, predacious, and mutualistic fungi (including mycorrhizae and lichen fungi). Many plant and animal parasites, and some mycorrhizal fungi, are facultative or interim saprotrophs that are able to persist in soil or decaying organic matter in the absence of a suitable host. *See* BIODEGRADATION; FUNGAL GENETICS; MYCORRHIZAE.

Communities. Communities and assemblages of fungi occur in all terrestrial and many aquatic and marine habitats. Regardless of domain, species composition and abundance are closely regulated by environment. Because fungi are heterotrophs, they are nutritionally dependent on the plants, animals, and microorganisms with which they coexist. Thus, the composition of fungal communities is strongly influenced by other biota. In the various biomes of the world, the fungal communities are as different and distinctive as are the plants and animals of those biomes. In temperate zone biomes, high species diversities for plants and microfungi are correlated, but the total species of fungi far exceeds that of plants.

As microorganisms, fungal communities can vary considerably in scale, from the surface inhabitants of a single leaf to a grassland soil. Individuals and species in fungal communities are functionally interconnected through their interactions. The ecological principles that govern distribution, abundance, and dynamics of species of plants and animals apply equally to fungal communities. Resource limitation, competition, niche specialization, trophic interactions (predation and parasitism), and antagonistic interactions all operate and affect species distribution and abundance in fungal communities. In fungal communities, as in communities of other organisms, disturbance can have a cascading effect leading to extinctions, changes in abundance, and appearance of previously undetected species.

Functions. Fungi interact with all organisms in ecosystems, directly or indirectly, and are key components in ecosystem processes. As decomposers, fungi are crucial in the process of nutrient cycling, including carbon cycling as well as the mineralization or immobilization of other elemental constituents. As parasites, pathogens, predators, mutualists, or food sources, fungi can directly influence the species composition and population dynamics of other organisms with which they coexist. Fungi may act both as agents of successional change or as factors contributing to resilience and stability. Introduced fungal pathogens have had a profound effect on the species composition of forests worldwide during the past century. Nitrogen-fixing lichens are important components of the nitrogen cycle. Fungi also form an essential link interconnecting functional components of food webs in many habitats.

Mycorrhizal fungi function as an interface between plant and soil, and are essential to the survival of most plants in natural habitats. Due to their fine hyphae and ability to permeate the substrate, mycorrhizal fungi enable their plant symbionts to extract nutrients from soil more efficiently. Mycorrhizal fungi extract essential nutrients from the inorganic pool, which are translocated to the plant symbiont. Generally, mycorrhizal plants are more vigorous than nonmycorrhizal plants. Approximately 70–90% of the vascular plant species worldwide are mycorrhizal in their natural habitats. Most mycorrhizal fungi have limited ability to decompose organic matter and rely on their plant symbionts for carbon nutrition, although some mycorrhizal fungi accomplish both decomposition and direct nutrient transfer.

Mycorrhizal fungi also link individuals and permit complex species interactions in ecosystems. The achlorophyllous vascular plant *Monotropa* obtains nutrients by digesting mycorrhizal hyphae that in turn receive nourishment from their tree associates. In the old-growth forests of the Pacific Northwest, truffles (fungi that form belowground sporocarps) in the genus *Hysterangium* are mycorrhizal with Douglas-fir. Most truffles produce odors attractive to rodents, which dig them up, eat them, and aid in their dispersal. *Hysterangium* and other mycorrhizal truffles are an important component of the diet of the northern flying squirrel, which in turn is preyed on by the northern spotted owl. A similar interdependence is known from the southwestern United States between ponderosa pine, the tassel-eared squirrel, and the mycorrhizal truffle *Rhizopogon*. Thus, the fungi form a crucial link in complex food webs.

Root pathogens of forest trees, such as *Phellinus* and *Armillaria*, can occur as large infection centers that expand radially, spreading by root contact and killing trees on the outer perimeter of the infestation. These infection centers contribute to age structure and heterogeneity in forests. Dying trees within the infection center lead to the formation of forest gaps where different, more disease-resistant tree species invade. As the infection centers age and expand, the trees nearer the center grow larger and a typical successional development occurs. Infection centers can be several hundred meters in diameter and over 1000 years old. Fungal plant pathogens, whether native or introduced, can profoundly affect the composition of the dominant vegetation. *See* HERBIVORY.

Applied ecology. There are several economically important areas that benefit from application of the principles of fungal ecology: biotechnology, biological control, bioremediation, agriculture, forestry, and land reclamation. Plant pathologists employ the principles of fungal ecology to study disease epidemics and control fungal pathogens in agricultural systems. Fungal biological control agents are used against weeds, pathogenic fungi, and insect pests. Northern jointvetch in the southeastern United States is successfully controlled by a *Colletotrichum*; *Heterobasidion* root disease of conifers is prevented from infecting cut stumps by the fungus *Phanerochaete gigantea*; the insect pathogen *Entomophaga maimaga* shows promise for control of the gypsy moth in the northeastern United States. Many fungi that attack insects are host-specific and thus can be used against specific pests without harming nontarget species.

Inoculation of tree seedlings for reforestation and landscape use with mycorrhizal fungi greatly improves their early establishment and survival. The increasing use of mycorrhizal fungi in agricultural production helps increase crop yields with less fertilization. The mycorrhizal fungus *Pisolithus* greatly improves survival of pines planted for reclamation of mine spoils. The ability of fungi to degrade complex substrates and tolerate high levels of toxic elements has led to their utilization for bioremediation of contaminated soil.

The observation that fungi produce substances that inhibit other microorganisms has led to discoveries of numerous antibiotics and other important pharmaceuticals, such as penicillin and cephalosporin from *Penicillium* and *Emericellopsis*, and cyclosporin from *Tolypocladium*. Mycotoxins such as aflatoxin in foodstuff are a human health concern. Fungal enzymes are used in numerous industrial processes; citric, itaconic, gallic, and gluconic acids are produced by *Aspergillus*; and several foods are produced through fungal fermentations. With only a small fraction of the total species known, the fungi offer a rich potential for bioprospecting, the search for novel genetic resources with unique, useful biochemical properties. *See* ECOLOGY; ECOSYSTEM; FUNGI; FUNGAL BIOTECHNOLOGY.

Martha Christensen; Jeffrey K. Stone

Ruminal gastrointestinal tract. A system of four chambers makes up the foregut of the ruminants. The rumen, reticulum, and omasum are pregastric chambers that are formed from modifications of the esophagus and the stomach. The rumen and reticulum together form a large fermentation vessel that is approximately 10 liters (3 gal) in sheep and 100–150 liters (26–40 gal) in cows. The rumen, which is the largest pregastric chamber, is a highly specialized environment. It contains heterogeneous plant material and microorganisms in a fluid amorphous and dynamic condition similar to a continuous culture chamber. Stratification may occur in the rumen as a result of particle size and specific gravity; a gas space exists above the partially digested feed. The pH of the rumen fluctuates between 5.8 and 6.8, depending on the frequency of feeding and feed type. However, a copious amount of saliva, which constantly enters with the feed, contains bicarbonate and phosphate buffers to maintain the pH. As a result of the animal's metabolism and heat from microbial fermentation, the temperature of the rumen is maintained between 38 and 40°C (100 and 104°F), depending on frequency of feed or water intake. The composition of the gases in the rumen is approximately 65% carbon dioxide (CO_2), 35% methane (CH_4), 7% nitrogen (N_2), 0.6% oxygen (O_2), 0.2% hydrogen (H_2), and 0.1% hydrogen sulfide (H_2S). The redox potential of the rumen ranges from -250 to -450 mV because facultative anaerobic microbes such as bacteria and yeast quickly utilize air that accompanies the feed into the rumen. *See* ANIMAL FEEDS.

Bacteria, protozoa, and fungi are the three major groups of microorganisms that inhabit the gastrointestinal tract. These microorganisms are suspended in the liquid phase, or attached to plant fragments, or attached to the lining of the rumen epithelium. Of the three groups of microbes, the fungi were most recently discovered. Three types of fungi occur in the gastrointestinal tract of herbivores. The first group includes facultatively anaerobic and aerobic fungi that are described as transient because they continually enter the rumen in feed and may not be able to grow under the anaerobic conditions of the rumen environment. The second group consists of two fungal species that parasitize ciliate protozoa. The third group consists of the obligately anaerobic zoosporic fungi that are saprophytic on plant material.

Interactions of fungi and microorganisms. Due to the complex nature of the rumen ecosystem, interactions between the different types of microorganisms occur. Interactions are known to occur between anaerobic fungi and methane-producing (methanogenic) bacteria, between anaerobic fungi and nonmethanogenic bacteria, and between anaerobic fungi and protozoa.

Methanogens. Anaerobic rumen fungi have synergistic relationships with methanogens. Cultures of an anaerobic fungus, *Neocallimastix frontalis*, degrade filter paper more extensively in the presence of methanogens than in their absence. Fermentation shifts from the formation of electron sink products such as lactate and hydrogen toward more reduced end products such as carbon dioxide and methane and acetate in the coculture. *Neocallimastix frontalis* alone causes 53% dry-matter loss in vitro, while cocultures of the fungus with *Methanosarcina barkeri* and *Methanobrevibacter* sp. lead to losses of 69% and 87%, respectively. A triculture of all three organisms results in a 98% loss of filter-paper dry weight, with the end products being carbon dioxide and methane.

Methanogens attach to the surfaces of the zoospore or thalli of the fungi when grown in coculture with fungi. Hydrogenosomes, enzymes involved in the production of hydrogen and carbon dioxide in anaerobic fungi, are located near the surface of the plasma membrane. It has been suggested that methanogens are able to readily utilize the products of the hydrogenosomes for growth. When *Piromyces* sp. and *Caecomyces* sp. are grown in cocultures with methanogens, there is an increase in acetate production. *See* METHANOGENESIS (BACTERIA).

Nonmethanogenic bacteria. Interactions between nonmethanogenic bacteria and anaerobic zoosporic fungi may be synergistic, symbiotic, or competitive. An example of a synergistic relationship is the growth of *Bacteroides ruminicola* in coculture with *N. frontalis* because degradation of xylan in the coculture is more than threefold greater than in monocultures of either the bacterium or fungus. *Selenomonas ruminantium*, a bacterium that consumes hydrogen, succinate, and lactate, increases cellulose digestion by rumen fungi. *Streptococcus bovis* and *Lachnospira multiparus* suppress xylan utilization by *N. frontalis*.

Protozoa. A two- to fourfold increase in zoospores and zoosporangia of anaerobic fungi in sheep is observed when protozoa are removed. Defaunation may be influenced by an animal's diet, since greater populations of zoospores are found in animals fed high-fiber diets than in those fed low-fiber or high-concentrate diets. The interaction between anaerobic fungi and rumen protozoa may be complementary rather than competitive in nature. *See* MICROBIAL ECOLOGY.

Daniel A. Wubah

Bibliography. L. R. Batra, *Insect-Fungus Symbiosis*, John Wiley, New York, 1979; G. C. Carroll and D. T. Wicklow, *The Fungal Community*, 2d ed., Marcel Dekker, New York, 1992; M. Christensen, A view of fungal ecology, *Mycologia*, 81:1–19, 1989; R. C. Cooke and A. D. M. Rayner, *The Ecology of Saprotrophic Fungi*, Cambridge University Press, 1984; N. J. Dix and J. Webster, *Fungal Ecology*, Chapman and Hall, London, 1995; S. Isaac, *Fungal Plant Interactions*, Chapman and Hall, London, 1992; D. A. Wubah, D. E. Akin, and W. S. Borneman, Biology, fiber degradation and enzymology of the anaerobic zoosporic fungi, *Crit. Rev. Microbiol.*, 19:99–115, 1993.

Fungal genetics

The study of gene structure and function in fungi. Genetic research has provided important knowledge about genes, heredity, genetic mechanisms, metabolism, physiology, and development in fungi, and in higher organisms in general, because in

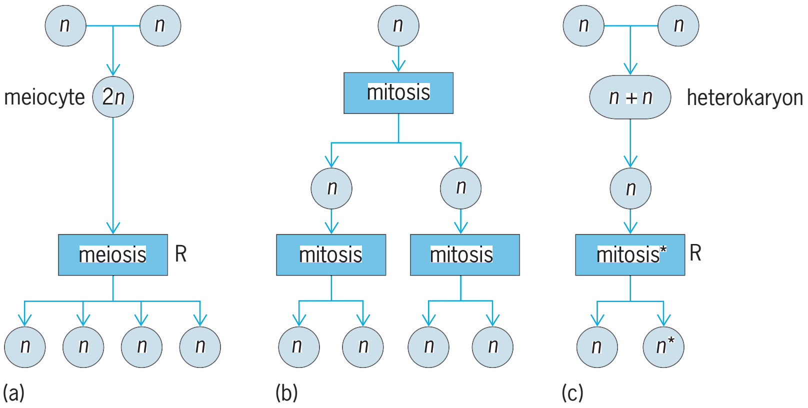


Fig. 1. Three different kinds of reproduction occurring in fungi, each of which provides opportunities for genetic analysis. (*a*) Sexual reproduction leads to recombination (R) of genes at meiosis. (*b*) Asexual reproduction (shown here in a typical haploid fungal cell) usually reproduces the gene set faithfully. (*c*) Parasexual reproduction derives from an atypical mitotic division of an unstable cell that produces haploid cells and other aneuploid unstable intermediates.

certain respects the fungal life cycle and cellular attributes are ideally suited to both mendelian and molecular genetic analysis. Fungal nuclei are predominantly haploid; that is, they contain only one set of chromosomes. This characteristic is useful in the study of mutations, which are usually recessive and therefore masked in diploid organisms. Mutational dissection is an important technique for the study of biological processes, and the use of haploid organisms conveniently allows for the immediate expression of mutant genes.

Reproduction in fungi can be asexual, sexual, or parasexual (**Fig. 1**). Asexual reproduction involves mitotic nuclear division during the growth of hyphae, cell division, or the production of asexual spores. Sexual reproduction is based on meiotic nuclear divisions fairly typical of eukaryotes in general. In ascomycetes and basidiomycetes, the spores, containing nuclei that are the four products of a single meiosis, remain together in a group called a tetrad. The isolation and testing of the phenotypes of cultures arising from the members of a tetrad (tetrad analysis) permit the study of the genetic events occurring in individual meioses; this possibility is offered by virtually no other eukaryotic group. In other groups, genetic analysis is limited to products recovered randomly from different meioses. Since a great deal of genetic analysis is based on meiosis, fungal tetrads have proved to be pivotal in shaping current ideas on this key process of eukaryotic biology. *See* EUKARYOTAE; MEIOSIS.

Because their preparation in large numbers is simple, fungal cells are useful in the study of rare events (such as mutations and recombinations) with frequencies as little as one in a million or less. In such cases, selective procedures must be used to identify cells derived from the rare events. The concepts and techniques of fungal asexual and parasexual genetics have been applied to the genetic manipulation of cultured cells of higher eukaryotes such as humans and green plants. However, the techniques remain much easier to perform with fungi.

Heterokaryons. In many fungi, notably the Fungi Imperfecti, there is no true sexual cycle based on meiosis. However, such fungi utilize a parasexual cycle which incorporates a unique reductive division more similar to mitosis than meiosis. The parasexual cycle uses the natural tendency of fungal cells to fuse, thus promoting nuclear and cytoplasmic mixing. The resulting mixed strains are called heterokaryons. Although the different nuclear types in a heterokaryon remain separate and haploid, rare nuclear fusions occur to create a transient diploid nucleus containing a chromosome set from each original strain. These diploid nuclei are unstable and promptly undergo atypical mitotic divisions leading to haploids and other unstable intermediate aneuploids. The derived haploid cells are commonly genetic recombinants. *See* DEUTEROMYCOTA; RECOMBINATION (GENETICS).

Heterokaryons are useful in other ways. They provide a convenient method of testing whether two recessive mutations are alleles by combining the two mutations in a heterokaryon. If the heterokaryon maintains the mutant phenotype of the two individual strains, the mutations can be assigned to the same gene; that is, they are alleles. Heterokaryons also provide a test for cytoplasmic inheritance. If heterokaryosis demonstrates that a novel phenotype is transmitted by cytoplasmic contact alone, then that phenotype must be based on cytoplasmic genes, probably those in the mitochondria. *See* ALLELE.

Genetics of mating type. The sexual cycle of heterothallic fungi requires fertilization by strains of different mating types. Generally, no outward difference is evident between these mating types. In bipolar species, mating type is determined by the alleles of one gene. For example, in *Saccharomyces cerevisiae* there are two alleles of the mating-type locus MAT, called a and α, whereas in the filamentous ascomycete *Neurospora crassa*, the alleles of the mt locus are called A and a. In tetrapolar species, such as many basidiomycetes, mating type is determined by two separate loci, A and B, and multiple alleles are found in nature. Sexual development occurs only if the mates differ for each locus: $A_1B_1 \times A_2B_2$ or $A_1B_2 \times A_2B_1$. $A_5B_1 \times A_1B_3$ are sexually compatible, but $A_5B_1 \times A_5B_3$ or $A_5B_3 \times A_1B_3$ are not.

In *Schizophyllum commune*, approximately 300 A and 100 B alleles are known. Recently the A loci were shown to contain pairs of genes encoding homeodomain proteins. The proteins derived from different A alleles form complexes within the fusion cells. These complexes are thought to bind to and regulate expression of genes essential for sexual development. The B loci have been shown to encode pheromones and serpentine transmembrane receptors. The types of molecules produced from both A and B loci are common types of regulators found in eukaryotic organisms. *See* BASIDIOMYCOTA; YEAST.

Mapping of genes and centromeres. In some ascomycetes, such as *Neurospora crassa*, the meiotic products in the tetrad (or octet, as in some fungi the postmeiotic nuclei soon undergo a mitosis) occur in a linear array within the ascus (*AAAAaaaa* or

AAaaAAaa). Thus, both the position of the meiotic products and the phenotypes which they display can be considered together. The frequencies of various phenotypes (*AB*, *Ab*, *aB*, or *ab*) and their positions within the octet allow estimates of the relative distance between two genes or between a gene and the contromere. *See* GENETIC MAPPING; MENDELISM.

Fungi lend themselves to gene mapping in two additional ways. Because fungal chromosomes are small, they may be separated by size when suspended in a gel matrix and then subjected to an electric field. Once the chromosomes have been separated by this pulsed-field gel electrophoresis, they may be visualized by staining, and subsequently transferred to a membrane by blotting. When this membrane is bathed with a solution containing a deoxyribonucleic acid (DNA) probe, made from a single, previously isolated gene, the probe binds to the chromosomes at the site of that gene. If the probe had been made radioactive, a photographic image of the blot can be visualized. In this way, specific genes may be mapped to particular chromosomes. Another way to map (and discover genes) is to determine the entire DNA sequence of the organism. Because fungi commonly have small genomes in comparison to other eukaryotes, it is feasible to sequence the entire genome. In fact, the first eukaryotic organism whose genome was completly sequenced was the baker's yeast, *S. cerevisiae.* The results demonstrated that yeast contains many genes that are also common to other eukaryotes, such as humans, but that the function of many genes remains unknown. This type of analysis has given rise to the new field of genomics, the study of genes and their functions.

Gene conversion and crossing-over. Within tetrads, occasional exceptions to Mendel's law of equal segregation of genes are seen; these rare exceptions have provided insight into the mechanism of crossing-over. The major exceptional types are *AAAAAaaa* (5:3 ratio) and *AAAAAAaa* (6:2 ratio). Both are examples of gene conversion, in which one or two copies of one allele have converted to the other allele, thus skewing in the 4:4 ratio. The regular observation of these types resulted in the heteroduplex theory of crossing-over, the prevailing theory for most organisms (**Fig. 2**). *See* CROSSING-OVER (GENETICS).

Mitochondrial inheritance. Inheritance determined by mitochondria may be suggested by uniparental inheritance or the heterokaryon test. The simplest type of uniparental inheritance is maternal, and one of the first examples was shown by the slow-growing phenotype, poky, in *Neurospora*. When a cross was made between a poky female and a normal male, all the progeny were poky, whereas in the reciprocal cross between a normal female and poky male, all the progeny were normal. The term female is used here, as commonly applied in fungal genetics, to identify the mate which contributes nearly all of the cytoplasm to the mating. Thus, poky was determined to be a cytoplasmically determined trait. Poky results from a mutation to the small circular DNA molecule carried within the mitochondrion.

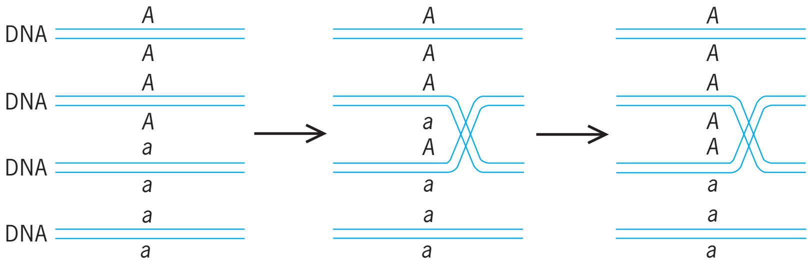
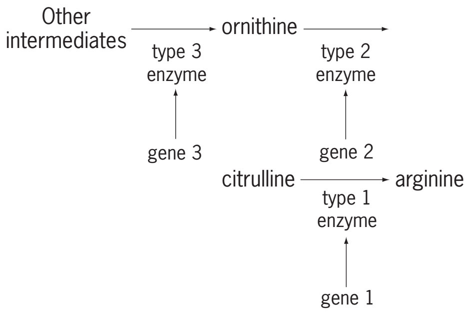


Fig. 2. Genetic crossing-over at the molecular level. As part of the process, heteroduplex DNA is formed (center). If the heteroduplex is corrected, as shown here from *a* to *A*, gene conversion is explained. Either or both heteroduplexes can be converted to *A* or to *a*, and hence a variety of aberrant ascus ratios are obtained; a 5:3 ratio is used here.

Study of fungal mitochondrial DNA has led to an understanding of the structure and function of these highly specialized organelles. Some surprises have resulted, including the discovery of a novel mitochondrial genetic code and a unique splicing mechanism of group I introns. *See* MITOCHONDRIA.

Genes and enzymes. The fact that each enzyme is coded by its own specific gene was first recognized in fungi and was of paramount importance because it showed how the many chemical reactions that take place in a living cell could be controlled by the genetic apparatus. The discovery arose from a biochemical study of nutritional mutants in *Neurospora*. The mutants were selected on the basis of their need for arginine in the growth medium (normal strains make their own arginine). Some mutants (type 1, for example) had an absolute requirement for arginine, whereas others (type 2) could also respond to a related compound, citrulline. A third type (3) responded to either citrulline or another related compound, ornithine. Those results can be explained if the mutations are in genes that code for enzymes that catalyze the sequential conversion of intermediates, including ornithine and citrulline, into arginine, as in the reaction below.



The enzyme deficiencies could be circumvented only by adding compounds normally synthesized after the block in each case. *See* ENZYME.

Genetic transformation. In genetically transformed organisms, the genome has been modified by the addition of DNA, a key technique in genetic engineering. Fungi were the first cells in which this key technique was shown to be feasible in eukaryotic organisms. In fungi, the cell wall is temporarily removed; exogenous DNA is then taken up by cells and the cell wall is restored. The incorporation of DNA must be detected by a suitable novel genetic marker included on the assimilated molecule in order to distinguish transformed from nontransformed cells. The fate of the DNA inside the cell depends largely on the nature of the vector or carrier. Some vectors can insert randomly throughout the genome. Others can be directed to specific sites, either inactivating a gene for some purpose or replacing a resident gene with an engineered version present on the vector. A third kind of vector remains uninserted as an autonomously replicating plasmid. The ability to transform fungal cells has contributed greatly to understanding the genetics, physiology, and development of fungi. Furthermore, the technique has permitted the engineering of fungi with modified metabolic properties for making products of utility in industry. *See* GENETIC ENGINEERING.

Fungal transposons. A surprising development in the molecular biology of eukaryotes was the discovery of transposons, pieces of DNA that can move to new locations in the chromosomes. Although transposons were once known only in bacteria, they are now recognized in many eukaryotes. The transposons found in fungi mobilize by either of two processes: one type via a ribonucleic acid (RNA) intermediate that is subsequently reverse-transcribed to DNA, and the other type via DNA directly. In either case, a DNA copy of the transposed segment is inserted into the new site and may contain, in addition to the transposon itself, segments of contagious DNA mobilized from the original chromosomal site. Because of the rearrangements which transposons may produce, they have been important in the evolution of the eukaryotic genome. *See* FUNGI; GENETICS; TRANSPOSONS.

A. J. F. Griffiths; R. Ullrich

Bibliography. J. W. Bennett and L. L. Lasure (eds.), *Gene Manipulations in Fungi*, 1985; J. W. Bennett and L. L. Lasure (eds.), *More Gene Manipulations in Fungi*, 1991; J. H. Burnett, *Mycogenetics*, 1975; J. R. S. Finchman, P. R. Day, and A. Radford, *Fungal Genetics*, 4th ed., 1979.

Fungal genomics

The study of the entire genetic content of organisms. It is a new discipline, developed out of advances in automated nucleic acid sequencing and computer technology. In the case of fungal genomics, the objective of the deoxyribonucleic acid (DNA) sequencing and analysis is the determination of the genomes of species in the kingdom Fungi. Fungal genomics seeks to establish the set of genes that makes a fungus specifically a fungus, as well as to determine what these species have in common with other organisms. Current genomics projects try to determine the entire genetic makeup of a given species by sequencing many overlapping fragments of DNA (at least 100,000 for a fungus), piecing these fragments together by computer, looking at the consensus sequence to identify genes and other regions of interest, and attempting to work out what these genes might do. Ultimately, this information is integrated into models of transcription (transcriptomics), protein expression (proteomics), and metabolic pathways (metabolomics). Sometimes the transcription, protein expression, and metabolic modeling stages of a genome project are called functional genomics. The deciphering of the resultant information has engaged computer scientists, mathematicians, and statisticians as well as biologists and chemists in a new discipline that is called bioinformatics. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID.

Fungal kingdom. The fungi are eukaryotic organisms that were once classified as plants but are now usually placed in their own kingdom. Four divisions (phyla) are distinguished within the fungal kingdom: Chytridiomycota, Zygomycota, Ascomycota, and Basidiomycota. In addition, fungi lacking a known sexual cycle are classified as Deuteromycota (Fungi Imperfecti). Taxonomists have reclassified the Oomycota, once called the lower fungi, into the kingdom Stramenopila. These so-called lower fungi include well-known species such as the causative agent of potato blight (*Phytophthora infestans*) that historically have been studied by mycologists. Out of tradition, the study of oomycete genomics usually is included in discussions of fungal genomics.

All true fungi are characterized by spore formation, an absorptive (osmotrophic) mode of nutrition and, with the exception of the chytrids, the absence of a motile stage in their life cycles. The best-known fungi are macroscopic organisms easily recognized by their distinctive fruiting bodies, for example, mushrooms, shelf fungi, and truffles. The majority of fungal species, however, are microscopic. Their common names—blight, mold, rot, and smut—reflect their capacity for causing both biodeterioration and disease. Some fungi are single-celled, of which the yeasts are the best known. The budding yeast, *Saccharomyces cerevisiae*, and the fission yeast, *Schizosaccharomyces pombe*, have been the targets of extensive genome projects. However, most macroscopic and microscopic fungi are filamentous, composed of threadlike cells called hyphae. The hypha is a distinctive tubular structure that endows filamentous fungi with cellular and genetic dynamics not found in other eukaryotic kingdoms. *See* FUNGI.

Economic importance. Fungi have enormous economic impact. They not only are major recyclers of organic materials in ecological cycles but also are important plant and animal pathogens. All fungi secrete degradative enzymes and organic acids into the environment. Many also make biologically active small molecules, some of which are toxic (for example, mushroom poisons, mycotoxins). Others have pharmaceutical value and are produced by industrial-scale fermentations for use in the manufacture of drugs such as penicillin, cephalosporin, and cyclosporin. Fungi also play a vital role in food fermentation (bread, wine, certain cheeses, soy sauce). In general, laboratory models or economically important fungi have been selected for fungal genome projects.

Genomic analysis. A typical genome project seeks to determine the complete DNA structure of a given organism and to identify and map all of its genes.

Shotgun phase. Today most sequencing centers use some version of the "shotgun" method, in which the genomic DNA of an organism is fragmented into small pieces, usually using high-frequency sound waves (sonication), and cloned to make a genomic library. The DNA from random clones is isolated and its sequence determined. Many steps in this process are automated, involving robots and expensive machines called DNA sequencers. Assembly involves putting together these random sequenced stretches of DNA into contigs (groups of cloned DNA representing overlapping segments of a particular chromosomal region) by using computer programs that compare individual sequences and find overlaps. Then genome assemblers put the fragments together into larger pieces. For large genomes, some of the most powerful computer arrays in the world are required for these computationally intense tasks. During this random or whole-genome shotgun phase, the idea is to sequence every bit of the genome at least six to ten times. However, some genome regions cannot be sequenced for various technical reasons. Often because of time and money constraints, genome-sequencing projects do not go beyond the whole-genome shotgun phase.

Finishing phase. The final "finishing" phase requires human input to join together the contigs. This process involves generating additional sequences to cover gaps and sequencing regions of the genome that were not cloned into the genomic libraries in the first place. The assembly and finishing steps work best when genetic maps, good karyotypes (organized arrays of chromosome pairs from a single cell), and/or reference genomes are available. For example, assembled sequences may be positioned to chromosomes using in-situ hybridization techniques, which enable identification of particular nucleic acid sequences while they remain in their original location in the nucleus. The last, labor-intensive stages of producing a finished sequence consume a disproportionate amount of time and money compared with the earlier, highly automated phases. For example, a draft of the human genome was announced early in 2000, but the complete version was not released until April 2003. *See* GENETIC MAPPING.

Annotation. Vast amounts of data are generated through automated sequencing. The ultimate goal is to associate DNA sequences with their functional roles in the cell. The most basic function-prediction method is the automated annotation of sequence. Programs such as Artemis, Genefinder, Genmaker, Glimmer M, and Grail are used to predict open reading frames (that is, genes are found using uninterrupted series of DNA triplets coding for amino acids without any stop codons), splice junctions (where coding and noncoding DNA meet), and other informational DNA. Once extracted, predicted genes can be subjected to sequence similarity (homology) searches to find homologs in nucleic acid or protein sequence data, usually using a member of the Basic Local Alignment Search Tool (BLAST) family of programs. Defining protein function is complicated because each gene can encode several distinct proteins. In general, when new protein coding regions are annotated, about half the genes are completely unknown. *See* GENETIC CODE; NUCLEIC ACID.

Postgenomics. Even annotated sequence information is not useful unless it leads to experimental analyses. After whole genome sequences are available, attention shifts from genome content to protein functions and systems biology. Linking unknown genes to the corresponding biologic function is variously called postgenomics or functional genomics. Genome data can be used to produce DNA microarrays, in which thousands of gene segments are deposited on a glass slide or chip and then used to simultaneously analyze the expression levels of the corresponding genes. For example, fungal molecules of ribonucleic acid (RNA) can be isolated from different time points in the sporulation and sexual cycle, reverse-transcribed to form copy DNA (cDNA), labeled with fluorescent chromophores, and the cDNAs tested against genomic microarrays. These data then are analyzed for expression patterns.

Fungal genome projects. The number of fungal genome projects is growing quickly, so it is difficult to give a comprehensive listing of the species that have been sequenced or are in the process of being sequenced. However, some of the major projects are described below.

Model organisms. The first eukaryotic organism with a completed genome was the budding yeast *S. cerevisiae*, completed in 1996. The genome size is slightly larger than 12 megabasepairs, with approximately 6000 identified genes. (Fungal genomes tend to be in the 10–50-Mb range, larger than bacteria but much smaller than mammals, insects, and most green plants.) Among all eukaryotes, *S. cerevisiae* provides the most sophisticated model for how cells work. However, its nonfilamentous nature reduces the value of comparisons with filamentous fungi.

The first filamentous fungus to reach the whole-genome shotgun phase, with its genome data in the public domain, was *Phanerochaete chrysosporium*, an organism sometimes called the model white rot. It has been studied extensively for its role in the degradation of lignin (the complex carbohydrate that makes wood hard) and for its ability to degrade a number of toxic environmental contaminants such as trinitrotoluene (TNT), atrazine, and benzo[a]-pyrene. Two other genetic models have also been sequenced: *Neurospora crassa*, the pink bread mold, and *Aspergillus nidulans* (also known as *Emericella nidulans*), a green mold that was used to elucidate the parasexual cycle. *See* FUNGAL GENETICS.

Economically important species. In addition to model organisms, fungal genome projects have targeted species of economic importance, in particular pathogens that cause plant and animal diseases. Medically important species that are the focus of genome projects include *Candida albicans*, an opportunistic pathogen that is commonly involved in vaginal infections and is the causative agent of thrush (whitish oral lesions); *Cryptococcus neoformans*, causative agent of cryptococcosis, a mycosis (fungal infection) that commonly spreads to the nervous system; *Histoplasma capsulatum*, causative agent of the most common human pulmonary mycosis; and *Pneumocystis carinii*, an unusual fungus that causes pneumonia, especially in patients with acquired immune deficiency syndrome (AIDS). Finally, *Aspergillus fumigatus*, the main agent associated with the disease aspergillosis, which is the biggest cause of infectious death in bone marrow transplant patients, is the target of an international genome project.

Plant pathogens with genome projects underway include *Ashbya gossypii*, a cotton pathogen; *Ustilago maydis*, cause of corn smut disease; and *Magnaporthe grisea*, causative agent of rice blast. Industrially important species such as *Aspergillus oryzae* (used in soy sauce and sake fermentations), *A. niger* (producer of enzymes and citric acid), *A. terreus* (source of the class of drugs called statins, which reduce blood pressure), *Trichoderma reseii* (maker of enzymes that break down the complex plant carbohydrate cellulose), and *Penicillium chryosgenum* (penicillin producer) are also the focus of genome projects in various stages of completion.

Applications. Comparative genomics is the study of differences and similarities in the content and organization of genomes among organisms. Genomic data coupled with microarray technology facilitate the study of networks of genes and pathways. Entire fungal genomes may be compared for the presence or absence of homologous pathways. A number of proprietary genome projects are also being conducted by pharmaceutical and agricultural companies to isolate new targets for drug screening. When new targets are discovered, libraries of small molecules can be screened to look for their ability to perturb particular fungus-specific pathways, in the hope of finding agents with narrow antifungal activity. Further, genomic analysis can provide information on organisms that are not easily cultivated in the laboratory.

Because genomics provides a new way of finding commonalities and differences among organisms, it facilitates analysis of comparative overviews of biological systems. Genome data can be compared with the genomes of selected eukaryotic models (for example, the fruitfly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans*) and with the genomes of other fungi.

The challenge is to come to an understanding of complex traits such as the ability of organisms to cause disease and produce toxic molecules. In fungi, regulation is often tied to biological circuits for carbon metabolism, nitrogen metabolism, and pH sensing. All of these regulatory circuits are proving to be more complex and more nuanced than many researchers expected. *See* FUNGAL BIOTECHNOLOGY; MOLECULAR BIOLOGY.

J. W. Bennett

Bibliography. J. W. Bennett and J. Arnold, Genomics for fungi, in R. Howard and N. Gow (eds.), *The Mycota VIII: Biology of the Fungal Cell*, pp. 267–297, Springer, 2001; A. Kumar et al., An integrated approach for finding overlooked genes in yeast, *Nat. Biotech*, 20:58–63, 2002; R. A. Prade and H. J. Bohnert (eds.), *Genomics of Plants and Fungi*, Marcel Dekker, New York, 2003; C. W. Sensen (ed.), *Essentials of Genomics and Bioinformatics*, Wiley-VCH, Weinheim, 2002.

Fungal infections

Invasion of the body caused by pathogenic fungi. Fungi are microorganisms that fill many niches in nature. Several thousand have been described, but fewer than 100 are routinely associated with invasive diseases of humans. In recent years there has been an alarming increase in the number of exotic and common species implicated in diseases of humans, particularly in those who have acquired immune deficiency syndrome (AIDS) and those who are being treated with broad-spectrum antibiotics or immunosuppressive agents. In general, healthy humans have a very high level of natural immunity to fungi, and most fungal infections are mild and self-limiting. Intact skin and mucosal surfaces and a functional immune system serve as the primary barriers to colonization by these ubiquitous organisms, but these barriers are sometimes breached (**Table 1**).

Unlike viruses, protozoan parasites, and some bacterial species, fungi do not require human or animal tissues to perpetuate or preserve the species. Virtually all fungi that have been implicated in human disease are free-living in nature. However, there are exceptions, including various *Candida* spp., which are frequently found on mucosal surfaces of the body such as the mouth and vagina, and *Malassezia furfur*, which is usually found on skin surfaces that are rich in sebaceous glands. These organisms are often cultured from healthy tissues, but under certain conditions they cause disease. Only a handful of fungi cause significant disease in healthy individuals. Once established, these diseases can be classified according to the tissues that are initially colonized (**Table 2**).

Superficial mycoses. Four infections are classified in the superficial mycoses. Black piedra, caused by *Piedraia hortai*, and white piedra, caused by *Trichosporn beigleii*, are infections of the hair. The skin infections include tinea nigra, caused by *Exophiala werneckii*, and tinea versicolor, caused by *M. furfur*. Where the skin is involved, the infections are limited to the outermost layers of the stratum corneum; in the case of hairs, the infection is limited to the cuticle. In general, these infections cause no physical discomfort to the patient, and the disease is brought to the attention of the physician for cosmetic reasons.

Cutaneous mycoses. The cutaneous mycoses are caused by a homogeneous group of keratinophilic fungi termed the dermatophytes. Species within this group are capable of colonizing the integument and its appendages (the hair and the nails). In general, the infections are limited to the nonliving keratinized layers of skin, hair, and nails, but a variety of pathologic changes can occur depending on the etiologic agent,

TABLE 1. Types of Diseases caused by fungi

Type	Description
Allergies	Hypersensitivity states acquired through exposure to a particular substance (allergen), for example, hypersensitivity pneumonitis
Mycotoxicoses	Infections caused by accidental or recreational ingestion of toxic fungal metabolites, such as mushroom poisoning
Infections	Colonization and invasion of living tissue by fungi, for example the mycoses Primary infection—the result of direct entry of the fungus into healthy susceptible tissues or organs such as the skin and lung Secondary infection—colonization of fungi in active or healed cavitary lesions such as those of tuberculosis, histoplasmosis, or infarcts, such as "fungus ball" Opportunist infections—growth of fungi in tissues of organs of immunosuppressed or otherwise debilitated patients, for example candidiasis

site of infection, and immune status of the host. The diseases are collectively called the dermatophytoses, ringworms, or tineas. They account for most of the fungal infections of humans.

This group of fungi is classified into the genera *Trichophyton*, *Microsporum*, and *Epidermophyton*. They elicit several well-defined clinical syndromes collectively called the tineas. The same species may be involved in different expressions of disease depending on the anatomic site involved. Examples are tinea capitis (ringworm of the scalp), tinea manuum (ringworm of the hand), tinea corporis (ringworm of the glabrous skin), tinea cruris (jock itch or ringworm of the groin), tinea unguium (ringworm of the nails), and tinea pedis (athlete's foot or ringworm of the foot).

There appears to be a delicate balance between host and parasite in dermatophyte infections. Some of these fungi show an evolutionary development toward a parasitic existence. Those that have achieved a high level of coexistence with humans also exhibit a degree of specificity toward tissues that are colonized. As versatile as these fungi seem, they do not appear to invade organs other than the keratinized layers of skin, hair, and nails.

Subcutaneous mycoses. The subcutaneous mycoses include a wide spectrum of infections caused by a heterogeneous group of fungi. The infections are characterized by the development of lesions at sites of inoculation, commonly as a result of traumatic implantation of the etiologic agent. The infections initially involve the deeper layers of the dermis and subcutaneous tissues, but they eventually extend into the epidermis. The lesions usually remain localized or spread slowly by direct extension via the lymphatics, for example, subcutaneous sporotrichosis. The type of pathology evoked is a result of the interplay between host and pathogen (**Table 3**).

TABLE 2. Categories of fungal infections

Type	Description
Superficial mycoses	Infections limited to the outermost layers of the stratum corneum or the hair shaft
Cutaneous mycoses	Infections of the epidermis, fingernails, toenails, and cortex of the hairs
Subcutaneous mycoses	Infections involving the dermis, subcutaneous tissues, fascia, or bone
Systemic mycoses	Infections originating in the lungs, sometimes spreading to other tissues by hematogenous or direct extension

Systemic mycoses. The initial focus of the systemic mycoses is the lung. The vast majority of cases in healthy, immunologically competent individuals are asymptomatic or of short duration and resolve rapidly, accompanied in the host by a high degree of specific resistance. However, in immunosuppressed patients the infection can lead to life-threatening disease.

Histoplasmosis. This mycosis, caused by *Histoplasma capsulatum*, is the most common respiratory fungus infection in the world. The organism grows in soil with a high nitrogen content, in particular, soils contaminated with the fecal droppings of bats and birds (especially chickens and starlings). Bats are naturally infected but birds are not. This disease is sometimes referred to as spelunker's disease because epidemics have occurred in individuals who have explored bat-infested caves and disturbed contaminated soil, resulting in exposure to the fungus. The disease is distributed throughout temperate, tropical, and subtropical areas of the world and is endemic in the Ohio and Mississippi valley regions of the United States, and scattered areas of Mexico and Central and South America. It is a mold in nature and a yeast in tissues of infected patients.

Blastomycosis. This disease is caused by *Blastomyces dermatitidis*. The organism has been isolated from soil and from woody specimens taken from a beaver dam that was the focus of an epidemic. The geographic distribution is limited to portions of North America and Africa. In the United States, the infection overlaps that for histoplasmosis and includes the Ohio and Mississippi valley rivers and the Missouri River basin. The organism is a mold in nature and a yeast, characterized by a broad-based attachment to the budding cell, in infected tissue.

Paracoccidioidomycosis. This disease, caused by *Paracoccidioides brasiliensis*, has a high incidence in Brazil, with endemic foci in Venezuela and Colombia. Epidemiologic surveys indicate that men and women have the same degree of exposure (50:50), but clinical disease is seen more often in males than females

TABLE 3. Subcutaneous mycoses

Disease	Clinical disease	Etiologic agents
Sporotrichosis	Lymphocutaneous sporotrichosis is the disease most commonly associated with this fungus. The lesions start as a painless nodule at the site of injury, followed by other nodules that develop along the lymphatics that drain the site. They become fluctuant and ulcerate. The infection rarely extends beyond the regional lymphatics.	*Sporothrix schenkii*
Chromoblastomycosis	The most common form of chromoblastomycosis consists of warty, vegetative lesions which take on the appearance of a cauliflower floret.	Various dematiaceous (darkly pigmented) fungi belonging to the genera *Fonsecaea, Wangiella, Phialophora*, and *Cladosporium*
Mycetoma	Localized indolent, deforming, swollen lesions are composed of abscesses, and draining sinuses with the presence of "grains" are microcolonies of the fungus.	*Pseudallescheria boydii, Madurella grisea, Acremonium killense, Leptosphaeria senegalensis*, and so on
Zygomycosis	The disease can run a rapid acute course ending in death or become chronic.	Organisms belonging to the phylum Zygomycota including species of *Mucor, Absidia, Cunninghamella*, and so on

(90:10). The organism is a mold in nature and a multipolar budding yeast in infected tissues.

Coccidioidomycosis. Caused by *Coccioides immitis*, this is truly a disease of the New World. Natural infections occur only in North, Central, and South America. Areas of high endemicity include the central San Joaquin valley in California, the southwestern United States, and certain regions of Argentina. It is a mold in nature and a spherule (sporangium) in tissues of the infected host.

Cryptococcosis. The etiologic agent for this disease, *Cryptococcus neoformans*, has been recovered in large numbers from the excreta of pigeons and their roosting areas. However, the organism does not infect these birds. It is the most common cause of fungal meningitis. It is an encapsulated yeast in tissues and culture. The disease is worldwide in distribution and is frequently associated with patients who are immunosuppressed or debilitated, but many patients may have no significant underlying immune or metabolic defect. *See* FUNGI; MEDICAL MYCOLOGY.

George S. Kobayashi

Bibliography. L. Ajello and R. J. Hay, *Medical Mycology*, vol. 4 of *Topley and Wilson's Microbiology and Microbial Infections*, ed. by L. Collier, A. Balows, and M. Sussman, Arnold, London, 1998; K. J. Kwon-Chung and J. E. Bennett, *Medical Mycology*, Lea & Febiger, Philadelphia, 1992; P. R. Murray et al., *Medical Microbiology*, 3d ed.,C. V. Mosby, St. Louis, 1998.

Fungal virus

Any of the viruses that infect fungi (mycoviruses). In general these viruses are spheres of 30–45-nanometer diameter composed of multiple units of a single protein arranged in an icosahedral structure enclosing a genome of segmented double-stranded ribonucleic acid (dsRNA). Viruses are found in most species of fungi, where they usually multiply without apparent harm to the host. Most fungal viruses are confined to closely related species in which they are transmitted only through sexual or asexual spores to progeny or by fusion of fungal hyphae (filamentous cells). Some fungal strains are infected with multiple virus species.

Characteristics. There are exceptions to each of the above generalizations. One mycovirus has been described that is greater than 45 nm in diameter and similar in structure to the deoxyribonucleic acid (DNA) of an animal virus, herpesvirus. This implies that it would have genomes of DNA instead of dsRNA. Some mycoviruses have genomes of single-stranded RNA instead of dsRNA. In some fungal species, virtually every specimen tested has viruses, whereas in other species they are rarely found.

Early studies of fungal viruses focused on their potential for biological control of important plant pathogenic fungi. This was prompted by Michael Hollings's observation that mushrooms killed by a disease known as dieback contained abundant virus-like particles. Subsequently scientists in the drug industry examined the potential effect of virus infection on the secretion of drugs such as penicillin from the fungi which produced them. Although many viruses were found, none appeared to affect the secretion of drugs. However, studies of killer toxins in *Saccharomyces cerevisiae* and *Ustilago maydis* showed an association between the presence of dsRNA mycoviruses and toxin production. Although hundreds of virus-containing fungi have been reported, very few have been studied in significant detail. Three families of mycoviruses are recognized by the International Committee on Taxonomy of Viruses. The most thoroughly studied mycoviruses are in the family Totiviridae.

Totiviridae. This family of viruses has strains in which the entire genetic sequence is contained in one strand of dsRNA. The *S. cerevisiae* virus has only two genes. One gene codes for the polymerase protein responsible for the replication of the virion within the fungal cells, and the second gene codes for the protein that makes up the capsid (shell) of the virus. *Saccharomyces cerevisiae* and *U. maydis* viruses are indistinguishable in appearance by cryo-electron microscopy at a resolution of 26 angstroms (2.6 nm). Their capsids are uniquely arranged in 60 pairs of protein subunits in icosahedral symmetry.

Saccharomyces cerevisiae is the yeast of commerce, and *U. maydis* is the corn smut fungus responsible for the annual loss of at least 5% of the world corn crop. Both of their viruses have strains that produce killer proteins (KP strains) that are toxic to other members of the same fungal species. Their potential importance to the yeast fermentation industry and to agriculture may be substantial. One *U. maydis* virus toxin (P4) has been crystallized and reconstructed on the computer to a resolution of 1.9 A (0.19 nm). It has 109 amino acids and structural similarity to scorpion toxin. The P4 toxin blocks calcium channels in experiments with mammalian cells, and there is evidence that it may kill fungal cells by interfering with calcium channel activity. Other *U. maydis* and *S. cerevisiae* KP toxins are known, and their nucleic acid coding sequences have been defined.

Partitiviridae. This family is composed of viruses in which genomes are segmented and the critical viral genes (polymerase and capsid) occur in two or more capsids. One of these viruses, *Penicillium stoloniferum* virus S, has been found in three distinct genera of fungi, and most have been found in various molds or plant pathogens. Since they have not been reported to affect human health or agriculture, they have not been studied as thoroughly as members of the Totiviridae.

Hypoviridae. This family has definite economic importance. It contains a single genus, *Hypovirus*, whose numerous strains have specific dsRNA segments that severely reduce the ability of the host fungus to cause disease (hypovirulence). This fungus, *C. parasitica*, is responsible for the decline of the American and European chestnut trees, once an important forest species and a source of wood for houses, furniture, fences, and fires. During the first half of the twentieth century, chestnut blight killed the standing part of most chestnut trees; however, roots survived. As sprouts from these roots were produced, most were also killed by the fungus; however, in recent years stands of chestnut trees in Europe and the United States have begun to recover, and all are infected with strains of *C. parasitica* carrying the hypovirulent strains of the virus. The hypoviruses are unusual in that they do not have a capsid enclosing the dsRNA genome. It appears that the dsRNA may be enclosed in a type of vacuole.

Plant disease. The above viruses are those for which sufficient biological or molecular data have been accumulated to place them in a taxonomic classification. Many other fungal viruses have been studied, but their placement in the classification system or their relationship to specific biological phenomena, such as plant disease, is tentative. The most prominent of these are the mushroom viruses.

The mushroom viruses are strongly associated with a serious disease known as dieback that occurs sporadically in commercial mushrooms throughout the world. There are at least three partially defined mushroom viruses, ten dsRNA segments characterized by size, and three dsRNA segments completely sequenced. Yet it is not clear which capsids and dsRNAs constitute a specific virus since the diseased mushrooms almost always have a complex of more than one virus. Nine of the dsRNA are closely associated with diseased mushrooms, but one commonly appears in healthy mushrooms. Most investigators feel that the mycoviruses are the cause of the dieback disease, but Koch's rules of proof are yet to be demonstrated.

Other plant diseases, in which there is a tentative association of a mycovirus with disease severity, include scab of potatoes caused by *Rhizoctonia solani*, the take-all disease of wheat caused by *Gaeumannomyces graminis*, and Dutch elm disease caused by *Ophiostoma ulmi*. *See* FUNGI; MYCOLOGY; PLANT PATHOLOGY; VIRUS; VIRUS CLASSIFICATION.

R. F. Bozarth

Bibliography. B. Chen and D. L. Nuss, Infectious cDNA clone of hypovirus CHV1-Euro7: A comparative virology approach to investigate virus-mediated hypovirulence of the chestnut blight fungus *Cryphonectria parasitica*, *J. Virol.*, 73:985–992, 1999; S. A. Ghabrial, New developments in fungal virology, *Adv. Virus Res.*, 43:303–388, 1994; F. Gu et al., Structure and function of a virally encoded fungal toxin from *Ustilago maydis:* A fungal and mammalian Ca^{2+} channel inhibitor, *Structure*, 3:805–814.

Fungi

Nucleated, usually filamentous, sporebearing organisms devoid of chlorophyll; typically reproducing both sexually and asexually; living as parasites in plants, animals, or other fungi, or as saprobes on plant or animal remains, in aquatic, marine, terrestrial, or subaerial habitats. Yeasts, mildews, rusts, mushrooms, and truffles are examples of fungi.

Phylogeny

Fungal phylogeny traditionally has been based on morphology, and there have been many schemes, each supported by different morphological characters or interpretations. Nucleotide sequences of genes coding for ribosomes have been compared to infer phylogenies for many organisms, including fungi. The fungal phylogeny most strongly supported by molecular information is also consistent with many basic morphological and biochemical features, so fungal phylogenetics is now much simpler. Organisms in the kingdom Fungi are mostly haploid, use chitin as a structural cell-wall polysaccharide, and synthesize lysine by the alpha amino adipic acid pathway; and their body is made of branching filaments (hyphae). The fungi arose about 1 billion years ago along with plants (including green algae), animals plus choanoflagellates, red algae, and stramenopiles. Ribosomal comparison indicates that the closest relatives to the fungi are the animals plus choanoflagellates.

There are significant life-cycle differences between the early diverging fungal divisions. In both of the early groups, haploid individuals mate to produce zygotes that typically serve as resting spores; the diploid nuclei in these spores meiose before germination. In the terminal groups, haploid individuals mate but the nuclei do not fuse immediately; as a result, zygote formation is delayed, leaving pairs of unfused nuclei in the hyphae. When the zygote is made, it does not usually serve as a resistant body; instead meiosis occurs directly, followed by spore formation.

Ascomycetes are the most numerous fungi (75% of all described species), and include lichen-forming symbionts. The group has traditionally been divided into unicellular yeasts and allies with naked asci, and hyphal forms with protected asci. However, ribosomal gene sequences indicate that some traditional yeasts and allied forms diverged early (early ascomycetes), at about the time ascomycetes were diverging from basidiomycetes. Hyphal ascomycetes protect their asci with a variety of fruiting bodies; the earliest fruiting bodies may have been open cups (Discomycetes), while in more recent groups they are flask shaped (Pyrenomycetes and Loculoascomycetes) or are completely closed (Plectomycetes). Ascomycetes lacking sexual structures have been classified in the Fungi Imperfecti, but molecular comparisons now allow their integration with the ascomycetes. John W. Taylor

General Characteristics

The mycelium, generally the vegetative body of fungi, is extremely variable. Unicellular forms, thought to be primitive or derived, grade into restricted mycelial forms; in most species, however, the mycelium is extensive and capable of indefinite growth. Some are typically perennial though most are ephemeral. The mycelium may be nonseptate, that is, coenocytic, with myriad scattered nuclei lying in a common cytoplasm, or septate, with each cell containing one to a very few nuclei or an indefinite number of nuclei. Septa may be either perforate or solid. Cell walls are composed largely of chitinlike materials except in one group of aquatic forms that have cellulose walls. Most mycelia are white, but a wide variety of pigments can be synthesized by specific forms and may be secreted into the medium or deposited in cell walls and protoplasm. Mycelial consistency varies from loose, soft wefts of hyphae to compact, hardened masses that resemble leather. Each cell is usually able to regenerate the entire mycelium, and vegetative propagation commonly results from mechanical fragmentation of the mycelium.

Reserve foods are stored in various fungi as fats and glycogen (animal starch); volutin granules are the typical reserve material of yeasts.

Asexual reproduction. Asexual reproduction, propagation by specialized elements that originate without sexual fusion, occurs in most species and is extremely diverse. The most common and important means of asexual reproduction are unicellular or multicellular spores of various types that swim, fall, blow, or are forcibly discharged from the parent mycelium (**Fig. 1**). The more important types of spores are the following.

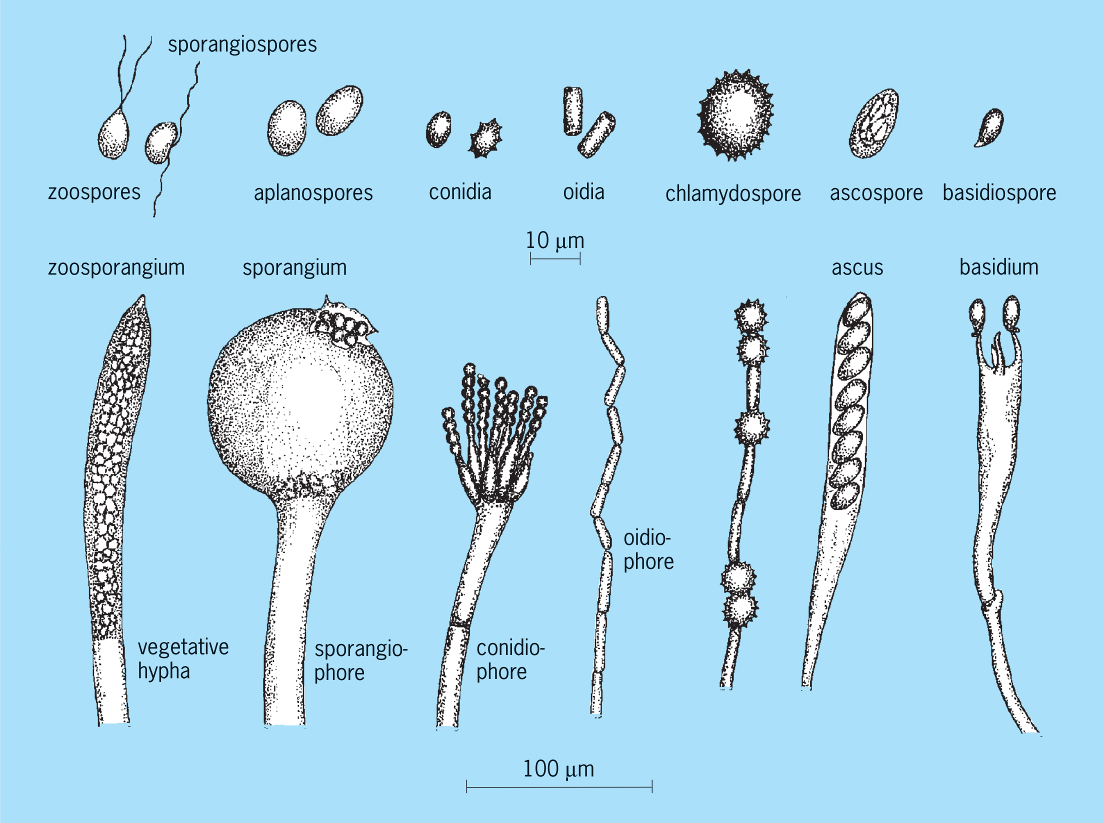


Fig. 1. Spore types and structure of parent mycelia in fungi.

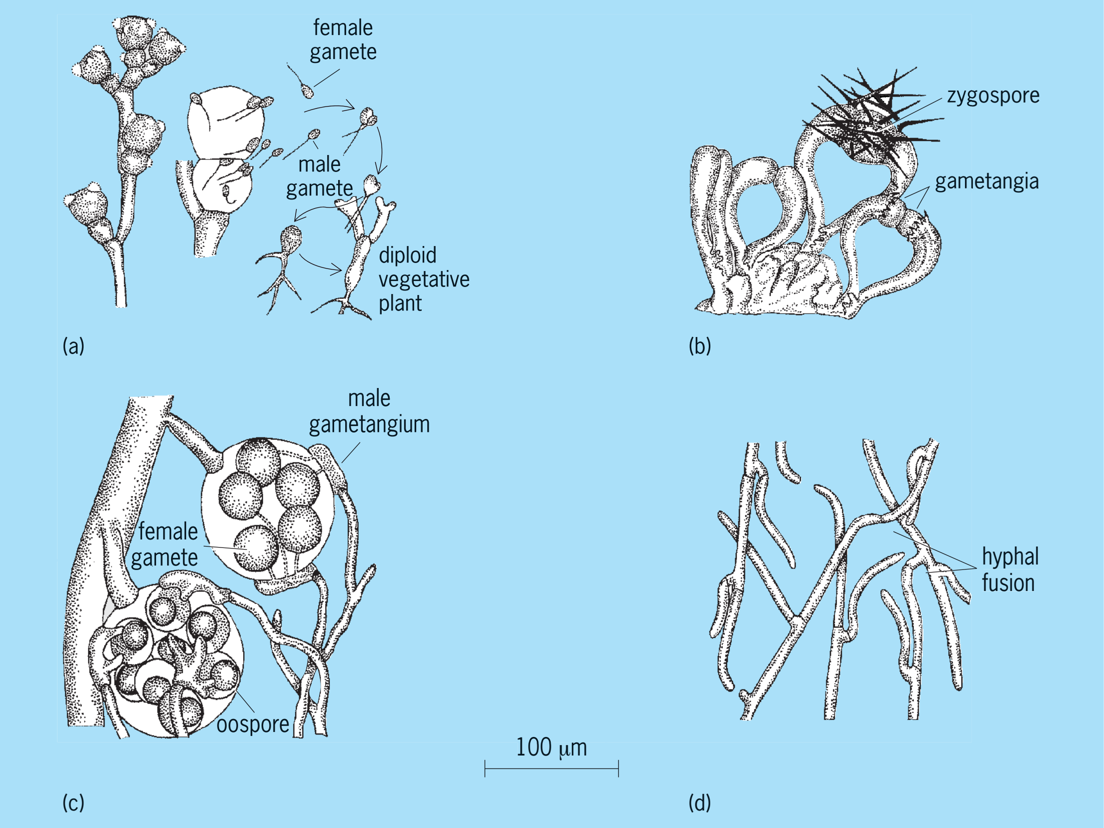


Fig. 2. Sexual mechanisms in fungi. (*a*) Gametic copulation. (*b*) Gametangial copulation. (*c*) Gamete-gametangial copulation. (*d*) Somatic copulation.

Sporangiospores are borne in unicellular sacs termed sporangia. Sporangia originate either by differentiation of vegetative cells or as more specialized, newly formed structures, frequently at the end of elongated stalks or sporangiophores. Sporangiospores are of two types: motile zoospores, equipped with flagella, and nonmotile aplanospores. The number of spores per sporangium varies from a few to thousands.

Conidia resemble nonmotile sporangiospores in shape, size, and structure but are produced externally upon a conidiophore which may be simple or quite elaborate.

Oidia, small, thin-walled cells that usually have flat ends, are produced by the autofragmentation of the vegetative hyphae.

Chlamydospores are thick-walled, nondeciduous spores interposed along vegetative hyphae.

Numerous other spore forms represent variants of the basic types above. Various pigments may be deposited in spore walls and account for mold colors in most cases. Asexual spores typically germinate by germ tubes that develop directly into vegetative mycelia.

Sexual reproduction. Sexual reproduction occurs in a majority of species of all classes. Juxtaposition and fusion of compatible sexual cells are achieved by four distinct sexual mechanisms:

Gametic copulation is the fusion in pairs of differentiated, uninucleate sexual cells or gametes formed in specialized sporangialike gametangia (**Fig. 2*a***). In isogamy, the two members of the fusion pair are alike; in anisogamy, they are morphologically different.

Gametangial copulation is the direct fusion of gametangia without actual differentation of gametes themselves (Fig. 2*b*).

Gamete-gametangial copulation is the fusion of a differentiated gamete of one sex with a gametangium of the other sex (Fig. 2*c*); the differentiated gamete may be either female or male.

Somatic copulation is the sexual fusion of undifferentiated vegetative cells (Fig. 2*d*).

Pattern of sexuality. This defines the underlying relationship between the two sexual elements that fuse in sexual reproduction; this relationship may be homothallism, which is fusion of elements of the same mycelium, or heterothallism, which is fusion only between elements of different mycelia. Obligatory cross-mating in heterothallic species is imposed by a number of distinct genetic devices that determine the mating competence of individuals with respect to sexual sign, that is, male and female, or to incompatibility class, that is, mating type. Incompatibility involves no morphological differentiation and in different groups may determine only two, hundreds, or even thousands of distinct self-sterile, but cross-fertile, mating types.

Product of sexual fusion. This and the time of nuclear fusion vary widely in the different classes and subgroups. Immediate nuclear fusion and the production of resting diploid spores, such as oospores and zygospores, or a diploid vegetative phase can also occur in lower forms. In more advanced forms, sexual fusion leads to the establishment of nuclear pairs, each a dikaryon, in hyphal systems capable of limited growth or indefinite growth. The dikaryotic mycelium can be parasitic within the closed, flask-shaped, or cup-shaped fruiting body known as the cleistothecium, perithecium, and apothecium, respectively, and its apical cells are differentiated as asci. Or, the dikaryotic mycelium can be capable of independent, indefinite vegetative growth and constitute the fruiting body of mushrooms and puffballs upon, or within, which the basidia are borne. Asci and basidia are saclike cells within which nuclear fusion and meiosis occur prior to production of endogenous ascospores and exogenous basidiospores, respectively.

Heterokaryosis. In heterokaryosis the association of genetically dissimilar nuclei in vegetative mycelia occurs.

Parasexuality. This is a vegetative process involving nuclear fusion and recombination of parental characters in the absence of sexual organs or fruiting structures. John R. Raper

Physiology

The functions and activities that fungi use to obtain food from the environment and the chemical modifications that they make to the materials taken up relates to fungal physiology. Fungi obtain organic substances (food) from their environment which have been produced through the (photosynthetic) activities of green plants, since fungi do not contain chlorophyll and are unable to manufacture their own food. Fungi are able to digest food externally by releasing enzymes into their environment which have the capacity to break down large molecules into smaller ones. These smaller molecules can be absorbed into the fungal body and transported to various locations where they can be used for energy or converted into different chemicals to make new cells or to serve other purposes. Some of the by-products of fungal metabolism may be useful to humans.

Fungi are multicellular organisms that obtain their food by absorption through cell walls. This characteristic distinguishes them from animals, which ingest their food, and from plants, which manufacture their food. Most fungi use nonliving plant material for food, but a few use nonliving animal material and therefore are called saprophytic organisms. In nature the decomposition of dead plant material is an important function of fungi, as the process releases nutrients back into the surrounding ecosystem where they can be reused by other organisms, including humans.

Decomposition. Fungal decomposition of plant material is very important in nature, as plant material contains high percentages of celluloses, lignins, and hemicelluloses which must be broken down to release the carbon, hydrogen, and oxygen. Fungi biodegrade the substrate, and as much as 70% of the dry organic matter is released as carbon dioxide and water through the process of respiration. Therefore, these fungi are highly involved in the carbon cycle of nature and influence the rate at which it proceeds. Fungi are the most effective decomposers of plant components, and evidence indicates that lignin decomposition is completed almost entirely by fungi. One gram of fertile soil may contain 10,000–100,000 fungal cells, many of which may be involved with different stages of the breakdown of plant material in the soil. *See* BIODEGRADATION; CELLULOSE; FUNGAL ECOLOGY; LIGNIN.

Pathogens. A few fungi have the physiological capability to grow on living plants and may cause diseases such as wheat rust or corn smut on these economically important plants. Some fungi can grow on grains and may produce substances known as aflatoxins which can be detrimental to animals or humans. A few species of fungi have the ability to grow and acquire their food from skin or hair on living animals such as cats, horses, and humans. The disease known as ringworm may result. It is not caused by a worm but by an expanding circular growth of a fungus which has the physiological capability to use the components of skin or hair as the food source. *See* AFLATOXIN; MEDICAL MYCOLOGY; PLANT PATHOLOGY.

The most frequently encountered fungal disease in humans is candidiasis, which is caused by one of the few fungi that is normally found associated with humans (*Candida albicans*). This organism does not ordinarily become a problem until the physiology of the body is changed or compromised in some way. Changes in hormonal level, general health, or immune status of the individual may alter the body physiology. The immune status may be altered by disease, prolonged broad-spectrum antibiotic treatment, immunosuppressant drugs associated with organ transplant, steroid therapy, or the prolonged use of drugs.

Industrial applications. Certain species of fungi have been used by humans since early times in the preparation of foods such as leavened bread, cheeses, and beverages. Additional by-products of fungal physiology are being used in industrial applications such as antibiotics, solvents, and pharmaceuticals. *See* FUNGAL BIOTECHNOLOGY; INDUSTRIAL MICROBIOLOGY; YEAST.

Nutritional requirements. While fungi share many common nutritional factors, groups or individual species of fungi may have specific requirements or capabilities to utilize different or unique food sources. The moisture level must be high for fungal growth to occur, and a relative humidity of about 100% usually supports maximum growth; however, most fungi do not require free water for growth. A few species can grow where relative humidity levels are as low as about 65%, but growth is usually not optimal.

Fungi usually require an organic source of carbon and a source of nitrogen which can generally be from an inorganic source. Carbon sources used by fungi are often carbohydrates, with the most common being sugars. Glucose is the sugar that most fungi have the ability to utilize at optimal rates. Minerals required by fungi are essentially the same ones that are required by green plants and other organisms.

Fungi are often able to make their needed vitamins; however, certain fungi may require an outside source of one or more vitamins. The need for a vitamin may be absolute (that is, growth is not possible without the vitamin) or partial, in which case, as more vitamin is added, the rate of growth increases. The most commonly reported vitamin deficiency in fungi is of the B-vitamin thiamine.

pH. The measure of the hydrogen-ion concentration in aqueous solutions is a measure of the acidity or alkalinity which may affect fungal growth. In general, fungi are able to grow over a relatively broad pH range, but a pH in the acidic range (below 7) is usually optimal. When growing in culture, fungi often change the pH of the medium. *See* PH.

Some fungi are facultative anaerobes, that is, they can live where there is no free oxygen, but can grow when oxygen is present. Most fungi are obligate aerobes and require some free molecular oxygen in their surroundings to support growth. Some species of fungi grown in shaken liquid cultures may have a growth rate as much as five times that of stationary cultures. Increased oxygenation of the culture medium seems to be a primary factor responsible for the increased growth.

The carbon dioxide concentration of air is approximately 0.03% (by volume); however, the concentration may increase because of cellular respiration in a wet soil or other confined environment. The concentration of carbon dioxide generally does not affect the growth of fungi until it becomes greater than the 10–15% range, and then it may become inhibitory. The concentration at which inhibition becomes measurable varies with different species of fungi.

Reproduction. Many fungi produce nonsexual spores as well as sexual spores in the reproductive process. Production of these spores is often affected by physiological process, including the nutritional status, presence of certain nutrients or metabolites, light, or the ratio of elements, such as carbon to nitrogen. The nutritional and environmental factors required for sporulation are often more specific than those required for vegetative growth, and may also demonstrate a linear relationship; that is, the greater the amount of the required factor, the greater the number of spores.

Physiological associations. Certain species of fungi produce chemicals that permit them to gain entrance to internal plant locations where they may grow and gain food from the host plant. Many of these fungi may become pathogenic and harm the host plant. A number of fungal species are able to enter plant roots and develop an association that may be beneficial to the plant under natural field conditions. This association of a higher plant root and a fungus that does not produce a disease is called a mycorrhiza. This fungal association with the plant root may permit the plant to live under soil conditions where it may not otherwise survive because of an excess of acid in the soil or a lack or excess of certain nutrients. While the presence or absence of this fungal association should be considered in the study of the nutritional status of plants, it is often overlooked. Consequently, under- or overestimation of the function of root systems has resulted. *See* MYCORRHIZAE.

Some species of fungi (belonging to the class Trichomycetes) are found growing within the intestinal tract of aquatic insect larvae (for example, the mosquito, midge, and blackfly) and other arthropods. Most of these species do not grow when they are removed from the intestinal tract and placed in a fungal food source. They have not been found growing elsewhere and, therefore, have an obligate association with the host, even though the host can do quite well without the fungus. They are apparently dependent on a nutritional component or combination of components most likely found only in the intestinal tract of the host. This specific physiological requirement is in contrast to other fungal species which are able to grow on a wide range of food sources. *See* FUNGAL GENETICS; FUNGAL VIRUS; FUNGISTAT AND FUNGICIDE; INSECT PATHOLOGY.

Marvin C. Williams

Bibliography. G. C. Ainsworth, F. K. Sparrow, and A. S. Sussman (eds.), *The Fungi*, vols. 7A and B, 1973; C. J. Alexopoulos and C. W. Mims, *Introductory Mycology*, 4th ed., 1995; L. Bobby (ed.), *Nitrogen, Phosphorous and Sulphur Utilization by Fungi*, 1989; T. D. Bruns, T. J. White, and J. W. Taylor, Fungal molecular systematics, *Annu. Rev. Ecol. Syst.*, 22:525–564, 1991; M. O. Garraway and R. C. Evans, *Fungal Nutrition and Physiology*, 1984, reprint 1991; D. L. Hawksworth, B. C. Sutton, and G. C. Ainsworth, *Ainsworth and Bisby's Dictionary of the Fungi*, 8th ed., 1996; B. Kendrick, *The Fifth Kingdom*, 2d ed., 1992; E. S. Luttrell, The function of taxonomy in mycology, *Mycologia*, 50:942–944, 1958; E. Moore-Landecker, *Fundamentals of the Fungi*, 4th ed., 1996; J. Webster, *Introduction to Fungi*, 2d ed., 1980.

Fungistat and fungicide

Synthetic or biosynthetic compounds used to control fungal diseases in animals and plants. A fungistat prevents the spread of a fungus, whereas a fungicide kills the fungus.

Agricultural fungicides. Chemical compounds are used to control plant diseases caused by fungi. Fungicides now used include both inorganic and organic compounds. Agricultural fungicides must have certain properties and conform to very strict regulations, and of thousands of compounds tested, few have reached the farmer's fields. Special equipment is needed to apply fungicides. *See* PLANT PATHOLOGY.

Requirements and regulations. The stipulations that must be met by manufacturers before they can sell fungicides and other pesticides have been drastically revised. In 1972 the Federal Environmental Protection and Control Act (FEPCA) substantially changed the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) of 1947. The FEPCA made the U.S. Environmental Protection Agency responsible for regulating the sale and use of pesticides within the country. The law states that all pesticides shall be classified according to their degree of toxicity to humans and other nontarget organisms: safer materials for "general use," and more toxic materials for "restricted use." All pesticides must still be proven effective before they may be registered for sale, and their labels must state all legal uses and the conditions of use. The law also requires that eventually each state must certify the qualifications of those who apply pesticides commercially. Fortunately, most agricultural fungicides are generally safer than the compounds used as insecticides or herbicides. *See* HERBICIDE; INSECTICIDE; PESTICIDE.

Because fungicides must now be labeled for specific uses on specific crops, the cost of development largely limits their use to such major crops as peanuts, citrus, potatoes, and apples. To fill the gap, the U.S. Department of Agriculture has established the IR-4 program to finance and facilitate label clearance for pesticide use on minor crops.

Formulation. The manner in which these compounds are applied, that is, as wettable powders, dusts, or emulsions, is often essential to the success of agricultural fungicides. Raw fungicides must be pulverized to uniform particles of the most effective size, mixed with wetting agents, or dissolved in solvents. These carriers or diluents must not degrade the fungicides or injure the plants.

Seed and soil treatments. Seeds and seedlings are protected against fungi in the soil by treating the seeds and the soil with fungicides. Seed-treating materials must be safe for seeds and must resist degradation by soil and soil microorganisms. Some soil fungicides are safe to use on living plants. An example is pentachloronitrobenzene, which can be drenched around seedlings of cruciferous crops and lettuce to protect them against root-rotting fungi. Other soil fungicides, such as formaldehyde, chloropicrin, and methyl isothiocyanate, are injurious to seeds and living plants. These compounds are useful because they are volatile. Used before planting, they have a chance to kill soil fungi and then escape from the soil.

Volatile fatty acid fungicides. Formic acid, acetic acid, and propionic acid up through pelargonic acid and capric acid (the C_1–C_{10} volatile fatty acids) possess significant fungicidal activity. Many of them are present in natural foodstuffs that are resistant to fungal attack. Use of the volatile fatty acids and their salts in bread to prevent ropy mold is widespread. A. I. Virtanen was awarded the Nobel prize in 1945 for his discovery and development of these lower volatile fatty acids to prevent fungal growth and so preserve the nutritious quality of cattle fodders. Formic acid, acetic acid, propionic acid, and isobutyric acid are used extensively for this purpose. Since these volatile fatty acids stop fungal growth, they prevent mycotoxin generation and lessen the risk of cancer from exposure to mycotoxins. *See* ANIMAL FEEDS.

Application methods. Dusters and sprayers are used to apply foliage fungicides. Conventional sprayers apply 300–500 gal/acre (2800–4700 liters/hectare) at pressures up to 600 lb/in.^{2} (4.1 megapascals). This equipment ensures the uniform, adequate coverage necessary for control. Among later developments in spray equipment are the mist blower and the low-pressure low-volume sprayer. The mist blower uses an air blast to spray droplets onto foliage. Mist blowers have been successful for applying fungicides to trees but are less satisfactory for applying fungicides to row crops. The low-pressure, low-volume sprayers are lightweight machines that apply about 80 gal of concentrated spray liquid per acre (240 liters per hectare) at a pressure of about 100 lb/in.^{2} (6.9 MPa). These have been successfully used to protect tomatoes and potatoes against diseases caused by fungi. A refinement of this method is ultra-low-volume (ULV) spraying. With ULV spraying, which employs special spinning cage micronizers, growers can protect certain crops by applying as little as 0.5–2 gal of spray liquid per acre (5–19 liters per hectare). Fungicides are also applied from aircraft.

James G. Horsfall; Saul Rich

Mode of action of fungicides. Two types of fungicides are used to control plant diseases: (1) Surface protectants remain on the plant surface and exert their toxic action on fungi before they have penetrated into plant tissue. (2) Systemic fungicides move into plant tissue and exert their toxic action on fungi which have already penetrated internally. These fungicides can also provide surface protection by acting on fungi before they have penetrated the plant.

Surface protectants. Prior to 1970, almost all fungicides used to control plant disease lacked the biochemical specificity necessary for therapeutic action within plant tissue; disease control by chemicals was limited essentially to surface protection. Before 1945, inorganic copper compounds such as Bordeaux mixture and copper carbonate and inorganic sulfur in the form of elemental sulfur and lime sulfur were the main fungicides available for plant protection. Copper is a multisite inhibitor that interferes with many different enzymes in fungal cells. Injury to plants is avoided by regulating the amount of soluble copper present on the plant surface. Elemental sulfur accepts electrons from the mitochondrial electron transport system and is reduced to hydrogen sulfide. This action interferes with energy generation and seems responsible for the fungitoxicity of sulfur.

After 1945, organic surface protectants became increasingly available for disease control. Among these fungicides are the dialkyldithiocarbamates, which include the fungicides thiram, ferbam, and ziram. These multisite inhibitors are capable of several kinds of toxic action in fungal cells, including metal chelation, mixed disulfide formation, and transport of heavy metals across membranes. Dialkyldithiocarbamates inhibit a multitude of enzymes, but the pyruvic dehydrogenase system is particularly sensitive to these fungicides.

The mode of action of ethylenebisdithiocarbamates, which include zineb, maneb, metiram, and mancozeb, differs from that of the dialkyldithiocarbamates. They undergo transformation to ethylenediisothiocyanate, which inactivates thiol groups of enzymes and metabolites in fungal cells. The phthalimides captan, folpet, and captafol and the nitrile fungicide chlorothalonil are important multisite inhibitors that also inactivate thiol groups of enzymes and metabolites in fungal cells.

The triphenyltins are surface protectants that interfere with oxidative phosphorylation in fungal mitochondria. They act at the same site as the antibiotic oligomycin.

Surface protectants are inexpensive and are rarely overwhelmed by fungal resistance. Therefore, they are still widely used for plant disease control even though many newer, more potent systemic fungicides have been developed.

Systemic fungicides. Most agricultural fungicides developed since 1965 are systemic compounds that act at a single target site in fungal cells. These fungicides comprise several mode-of-action groups.

The carboxamides, which include carboxin, oxycarboxin, mebenil, pyracarbolid, benodanil, mepronil, flutolanil, and furmecyclox, control basidiomycete-type fungi. Studies of carboxin indicate that the toxicity of the carboxamides is due to specific inhibition of succinic dehydrogenase.

The benzimidazoles benomyl, carbendazim, and thiabendazole and the phenylcarbamate diethofencarb specifically interfere with the formation of microtubules which function in a variety of cellular processes, including mitosis and maintenance of cell shape. These fungicides bind specifically to protein subunits called tubulin, and prevent their assembly to form microtubules.

The largest and most important group of systemic fungicides used to control fungal diseases of plants and humans are inhibitors of ergosterol and other desmethyl sterol biosynthesis. These fungicides are often referred to as ergosterol biosynthesis inhibitors or sterol biosynthesis inhibitors. The ergosterol biosynthesis inhibitors can be divided into three groups based on their specific site of action in the sterol biosynthetic pathway. (1) Squalene epoxidase inhibitors, which include the fungicides tolnaftate, naftifine, and terbinafine, have been developed for medical uses. (2) Inhibitors of the sterol C-14 demethylase include numerous fungicides (see **table**). (3) Inhibitors of the sterol $\Delta^8 \rightarrow \Delta^7$ isomerase, where the double bond is shifted from the 8(9) position to the 7(8) position, include the morpholine fungicides tridemorph, dodemorph, aldimorph, trimorphamide, and fenpropimorph. The morpholine fungicides also inhibit Δ^{14}-reductase, the enzyme that reduces the double bond at position 14(15) in the sterol structure to saturation, in some fungi. Morpholines are used to control plant fungal diseases but are not used for medicinal purposes. The action of ergosterol biosynthesis inhibitors results in a deficiency of the appropriate sterol (usually ergosterol) necessary for proper membrane structure and function and causes cell death or inhibition of growth.

Fungicides that inhibit sterol C-14 demethylase as their mode of action

Chemical class and name	Use*	Chemical class and name	Use*
Imidazole		*Triazole*	
Imazalil	P	Triadimefon	P
Fenapanil	P	Triadimenol	P
Prochloraz	P	Prociconazole	P
Miconazole	M	Etaconazole	P
Ketoconazole	M	Diniconazol	P
Clotrimazole	M	Bitertanol	P
Econazole	M	Diclobutrazol	P
Isoconazole	M	Flutraifol	P
		Fusilazol	P
Pyrimidine		Myclobutanil	P
		Penconazole	P
Fenarimol	P	Itraconazole	P
Nuarimol	P	Terconazole	M
Pyridine		*Piperazine*	
Buthiobate	P	Triforine	P
Pyrifenox			

*P = used or developed to control fungal disease of plants. M = used or developed for medicinal purposes.

The dicarboximide fungicides iprodione, procymidone, vinclozolin, chlozolinate, and metomeclan are particularly valuable for the control of plant diseases caused by species of *Botrytis*, *Sclerotinia*, *Monilinia*, *Alternaria*, *Sclerotium*, and *Phoma*. The modes of action of those compounds remain uncertain, although they are known to inhibit certain flavin enzymes. That action is believed to result in the production of reactive oxygen radicals that destroy membranes by peroxidation of fatty acids in phospholipids and attack nucleic acids and proteins. The dicarboximides are believed to act like the aromatic hydrocarbon fungicides chloroneb, quintozene, and dicloran because fungi that develop resistance to dicarboximides usually are resistant also to the aromatic hydrocarbons.

The phenylamide fungicides metalaxyl, furalaxyl, benalaxyl, ofurace, cyprofuram, and oxadixyl are used to control fungal diseases of plants. Their toxicity is limited primarily to fungi in the order Peronosporales and is due to selective inhibition of ribosomal ribonucleic acid synthesis.

Several antibiotics produced by microorganisms are used to control fungal diseases of plants or animals. The polyene antibiotics amphotericin B and nystatin are widely used to control fungal infections of humans. They bind to ergosterol in fungal membranes and disrupt selective membrane permeability, which leads to cell death.

Cycloheximide, a compound widely used as a specific inhibitor of protein synthesis in eukaryotic organisms, is a potent antifungal agent. Because of its toxicity to higher plants, it is used only to a limited extent for the control of fungal diseases of plants. The antibiotic inhibits protein synthesis by blocking polymerization of amino acids by fungal ribosomes.

The antibiotics blasticidin S and kasugamycin are also specific inhibitors of protein synthesis. They are used primarily in Japan to control rice blast and other fungal diseases of plants.

The polyoxin antibiotics inhibit chitin synthetase, leading to an interference with fungal cell wall synthesis and the distortion and rupture of germ tubes. Polyoxins are used in Japan to control rice blast and several fungal diseases of fruits and vegetables.

Most fungicides control disease by direct toxicity to the growth of fungal pathogens in their saprophytic as well as parasitic phases, but diseases can be controlled by inhibiting activities related specifically to parasitism. The melanin biosynthesis inhibitors tricyclazole and pyroquilon are not toxic to saprophytic growth of the rice blast fungus, *Pyricularia oryzae*, but block the fungus from penetrating the plant epidermis and prevent it from becoming established as a parasite. That is accomplished by inhibiting reactions in the polyketide pathway to melanin. This pathway is induced after the germinated fungus spores have formed appressoria (penetrating structures) on the plant epidermis. It remains unclear why blocking melanin biosynthesis prevents epidermal penetration by appressoria, but evidence points to decreased rigidity of appressorial walls, to a weakened adhesion of appressoria to the plant surface, or to an accumulation of toxic melanin precursors.

Resistance to agricultural fungicides. There have been very few instances of fungal resistance to surface protectant fungicides like inorganic copper compounds, dithiocarbamates, and phthalimides, even though they have been used for more than 35 years. The absence of resistance of target fungi exposed to those fungicides is attributed to their action at multiple sites within fungal cells. However, most systemic fungicides that act at a single site have generated serious problems with fungal resistance. A single gene mutation in a fungus can lead to loss of effectiveness of all fungicides in a particular mode-of-action group. Experience has shown that frequent, uninterrupted use of a fungicide increases the risk for development of a resistant strain of the target organism. That risk can be reduced by alternating use of fungicides with different modes of action or by using them in mixtures. The emergence of a fungicide resistance problem and its persistence in the absence of fungicide use are dependent on the fitness of the resistant strain. If the resistant strain is as fit as the wild-type strain, the former may persist with little reduction in the fungal propulation for many years, whether or not the fungicide is used. Such has often occurred with benzimidazole-resistant fungal strains. On the other hand, if fitness of the resistant strain is lower than that of the sensitive, wild-type strain, the resistant strain may decline to a low percentage of the population in one or two seasons, and the fungicide may again be effective for a limited number of treatments.

The mode of action of a systemic fungicide is also a factor in the risk for resistance development. Benzimidazoles and phenylamides are high-risk fungicides, whereas the ergosterol biosynthesis inhibitors are low- or moderate-risk fungicides. Most mutants resistant to ergosterol biosynthesis inhibitors appear to have a reduced fitness.

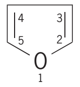
The principal mechanism of resistance is decreased affinity of the target sites for the fungicide. This mechanism of resistance has been observed for benzimidazoles, phenylamides, carboxamides, cycloheximide, and kasugamycin. In cases of resistance to benzimidazoles and carboxamides, new fungicides have been designed to fit the modified target sites, and these are highly toxic to the resistant strains. An example is the phenylcarbamate diethofencarb, which was designed to control benzimidazole-resistant strains of *Botrytis cinerea*; it is used in vineyards in France. *See* FUNGI.

Hugh D. Sisler

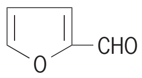
Bibliography. C. J. Delp (ed.), *Fungicide Resistance in North America*, 1988; Y. L. Nene and P. N. Thapliyal, *Fungicides in Plant Disease Control*, 3d ed., 1993; D. Pimentel (ed.), *Handbook of Pest Management in Agriculture*, 2d ed., 1990; M. R. Siegel and H. D. Sisler (eds.), *Antifungal Compounds*, 1977; A. J. P. Trinci and J. F. Ryley (eds.), *Mode of Action of Antifungal Agents*, 1984.

Furan

One of a group of organic heterocyclic compounds containing a diunsaturated ring of four carbon atoms and one oxygen atom. Furan (**1**) is a typical member of the group. Furfural (**2**) and some of its close relatives, such as furfuryl alcohol, tetrahydrofurfuryl alcohol, and tetrahydrofuran, are important chemicals of commerce. *See* HETEROCYCLIC COMPOUNDS.



(**1**)

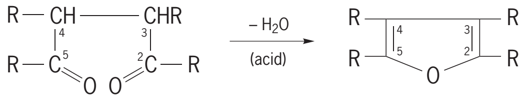


(**2**)

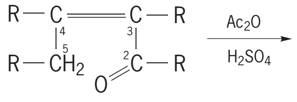
Properties and preparation. Furan (**1**) is a colorless, volatile liquid, bp 31.4°C (88.5°F), n_D^{20} 1.42140, density (20/4) 0.9378, which is stable to alkali but not to mineral acid. Its water solubility is approximately 1% at room temperature. On exposure to air, furan decomposes very slowly by autoxidation. Substituted furans, particularly negatively substituted furans, are much less sensitive. The furan system is aromatic, with resonance energies of 17–25 kcal/mole reported. Nitration, halogenation, acylation, mercuration, and sulfonation reactions occur with relative ease. The extreme sensitivity of furan itself to strong acid precludes many direct electrophilic substitutions. The presence of negative groups stabilizes the ring. Unoccupied 2 or 5 positions are favored as the site of substitution.

Addition reactions, in contradistinction to substitution reactions, have also been observed. Furan, for example, condenses with dienophiles in the Diels-Alder reaction. *See* DIELS-ALDER REACTION.

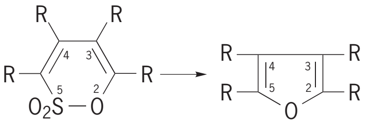
The most versatile general furan synthesis is the acid-catalyzed cyclization of 1,4-dicarbonyl compounds (**3**). The R groups in (**3**) may be hydrogen, aryl, alkyl, or carbethoxyl. Another synthesis (Treibs) transforms an α,β-unsaturated ketone (**4**) to an unsaturated δ-sulfone (**5**), which on pyrolysis loses sulfur dioxide and forms the furan. Furans are also formed in the acid-catalyzed dehydration of carbohydrates, the most important example being the production of furfural from the polysaccharides present in annually renewable agricultural waste materials (**6**). Furan is obtained from furfural either directly by catalytic decarbonylation or by oxidation to furoic acid (**7**) followed by decarboxylation.



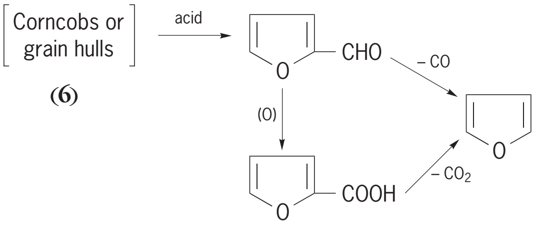
(**3**)



(**4**)



(**5**)

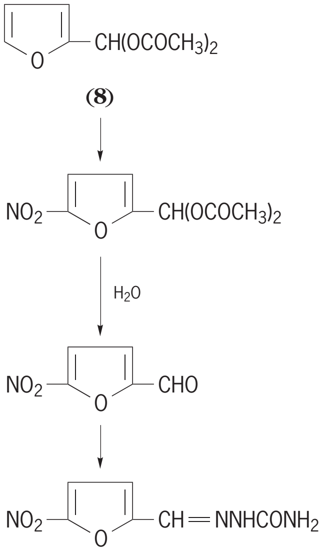


(**6**)

(**7**)

Important derivatives. Oxidation of furans generally disrupts the system, although some ring oxidations give useful products. Ring reduction is possible, the standard method being catalytic rather than chemical. Several products derived from furfural directly or indirectly by hydrogenation are of considerable industrial importance.

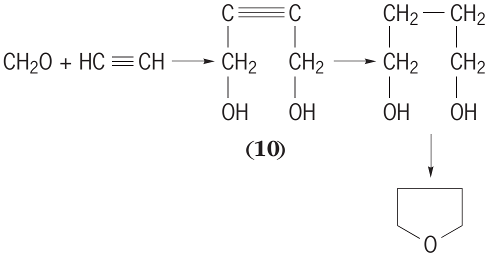
Reaction of furans with halogens, although not altogether free of complications, can be used for the synthesis of halofurans. Furansulfonic acids are prepared by sulfonation with sulfuric acid, or with a sulfur trioxide–pyridine complex. Nitration of furan compounds generally makes use of fuming nitric acid in acetic anhydride solvent. The synthesis of the bacteriostatic and bacteriocidal agent nitrofurazone, furacin (**9**), involves such a nitration of derivative (**8**).



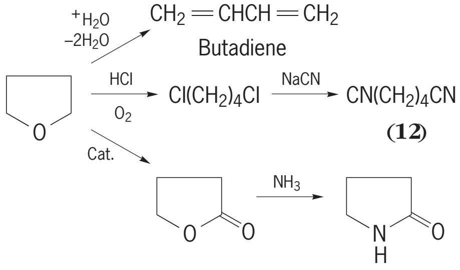
(**9**)

Furan aldehydes, prepared by formylation with hydrogen cyanide and hydrogen chloride, resemble benzaldehyde in their functional group reactions. Furyl ketones, prepared by Friedel-Crafts acylation, react normally. All the furan mono- and polycarboxylic acids are known. Removal of the carboxyl group from the 2 or 5 positions by decarboxylation is easier than from the 3 or 4 positions.

Tetrahydrofuran (**11**) is prepared either by decarbonylation of furfural to furan followed by catalytic hydrogenation, or by the Reppe process from butynediol (**10**). Tetrahydrofuran is a water-miscible, colorless liquid, bp 66°C (151°F), density (20/4) 0.8880, and n_D^{20} 1.4073. It is an industrial solvent, and also has served as an intermediate in the synthesis of butadiene for polymerization, adiponitrile (**12**) for nylon, and pyrrolidone (**13**) for the blood extender polyvinylpyrrolidone.



(**11**)



(**13**)

Walter J. Gensler

Bibliography. A. P. Dunlop and F. N. Peters, *The Furans*, Amer. Chem. Soc. Monogr. no. 119, 1953; G. P. Ellis, *Synthesis of Fused Heterocycles*, 1987; R. J.

Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; T. L. Gilchrist, *Heterocyclic Chemistry*, 3d ed., 1997.

Furnace construction

A furnace is an apparatus in which heat is liberated and transferred directly or indirectly to a solid or fluid mass for the purpose of effecting a physical or chemical change. The source of heat is the energy released in the oxidation of fossil fuel (commonly known as combustion) or the flow of electric current through adjacent semiconductors or through the mass to be heated. In recent years, scientific and engineering effort has been made to utilize nuclear and solar energy for heating purposes. Therefore, according to the source of heat and method of its application, there are four categories of furnaces; combustion, electric, nuclear, and solar, in the order of their present commercial or industrial importance.

Furnaces employing combustion vary widely in construction, depending upon the application of the heat released, whether direct or indirect. Direct heat transfer is used, in regenerative refractory-type heaters, in flow systems in which reactants are injected into the combustion gases. Indirect heat transfer is employed in heaters in which the mass to be heated is kept separate from the combustion gases and made to flow in tubes which absorb and transmit the heat to the fluid to be heated.

Furnaces developed for indirect heat transfer can be divided into two classes. One class of heaters is used solely for general utility purposes, such as all types of boilers; and the second class is applied in the petroleum and chemical industries as an essential unit operation in refining or processing plants. The second category of furnaces will be discussed in this article. *See* BOILER.

Directly fired furnaces are employed in oil refineries and chemical process units whenever the temperature level to which a fluid must be heated is above that attainable with utility steam. Furnaces as heat-transfer apparatuses generally cost more per thermal unit of heat transferred than conventional tubular heat exchangers. According to the kind of service, furnaces in process units can be divided into two classes:

1 Those which perform solely a heating duty, that is, raising the temperature of a fluid and effecting essentially no change of state or of chemical composition. Furnaces which perform this one duty may be termed conventional heaters.

2 Those which handle a fluid undergoing a change during heating. The physical or chemical changes constitute an essential performance requirement. Typical applications of this order are associated with such processes as distillation or preheating of temperature-sensitive materials, pyrolysis of hydrocarbons or organic chemicals, and catalytic steam-gas reforming for the production of synthesis gas.

The conventional tubular heater came into industrial use, especially in oil refineries, about 1925 and was termed a tube still to distinguish it from the shell still, a horizontal cylindrical vessel mounted on top of a firebox. The design of these tube stills or furnaces, which is essentially the same as originally conceived by analogy with developments in steam-generating equipment, consists of pipes connected by 180° return bends forming a continuous coil and arranged in a refractory furnace setting, partly in the combustion chamber. Heat is absorbed mainly by radiation and partly in a confined flue-gas passage from the combustion chamber in which heat is absorbed mainly by convection. The fluid flow through the coil is generally countercurrent to the flow of combustion gases, first through convection tubes and then through the radiant-tube section; thus a reasonable thermal efficiency can readily be attained by providing convection tube surface to an economically justifiable extent. *See* ELECTRIC FURNACE; FIRED HEATER; PETROLEUM PROCESSING AND REFINING; STEAM-GENERATING FURNACE.

The furnaces provided in modern process units vary considerably in outer shape, arrangement of tubes, and type and location of burners, depending mainly upon the desire of the designer or manufacturer to be identified with a specific model. Some distinctive furnace designs are shown in **Fig. 1**.

Design. The furnace design for a given performance or thermal efficiency is usually evolved by the following procedure: (1) determination of the composition of the combustion products and the amount of the liberated heat which must be utilized to meet the postulated thermal efficiency; (2) allocation of heat to be absorbed by the heating elements located in the combustion or radiant chamber and in the convection section; (3) determination of the heat-transfer rate and heating surface area in the radiant section; and (4) determination of the heat-transfer rate and tube surface area in the convection section or sections.

Source of heat. The combustion products vary in composition according to the type of fossil fuel burned and the excess air used in the oxidation process. The carbon and hydrocarbon content of the fuel governs the combustion-gas composition, and many formulas relating it to the elemental, oxidizable constituents of the fuel have appeared in the literature. Complete combustion is a prerequisite for high thermal efficiency; to ensure this, air is used in excess of the minimum or stoichiometric requirements, depending upon the type of fuel and the combustion equipment or system (**Table 1**).

The type of fuel, its heating value, and the excess air applied in combustion determine the theoretical flame temperature which would prevail if the oxidation were instantaneous. However, combustion is a rate process requiring time, and the actual flame temperature is considerably lower because of the radiation from the combustion zone. The theoretical flame temperature can readily be determined from the heating value of the fuel and the amount of combustion products evolved, including the excess air. Accurate enthalpy data on the combustion gases are

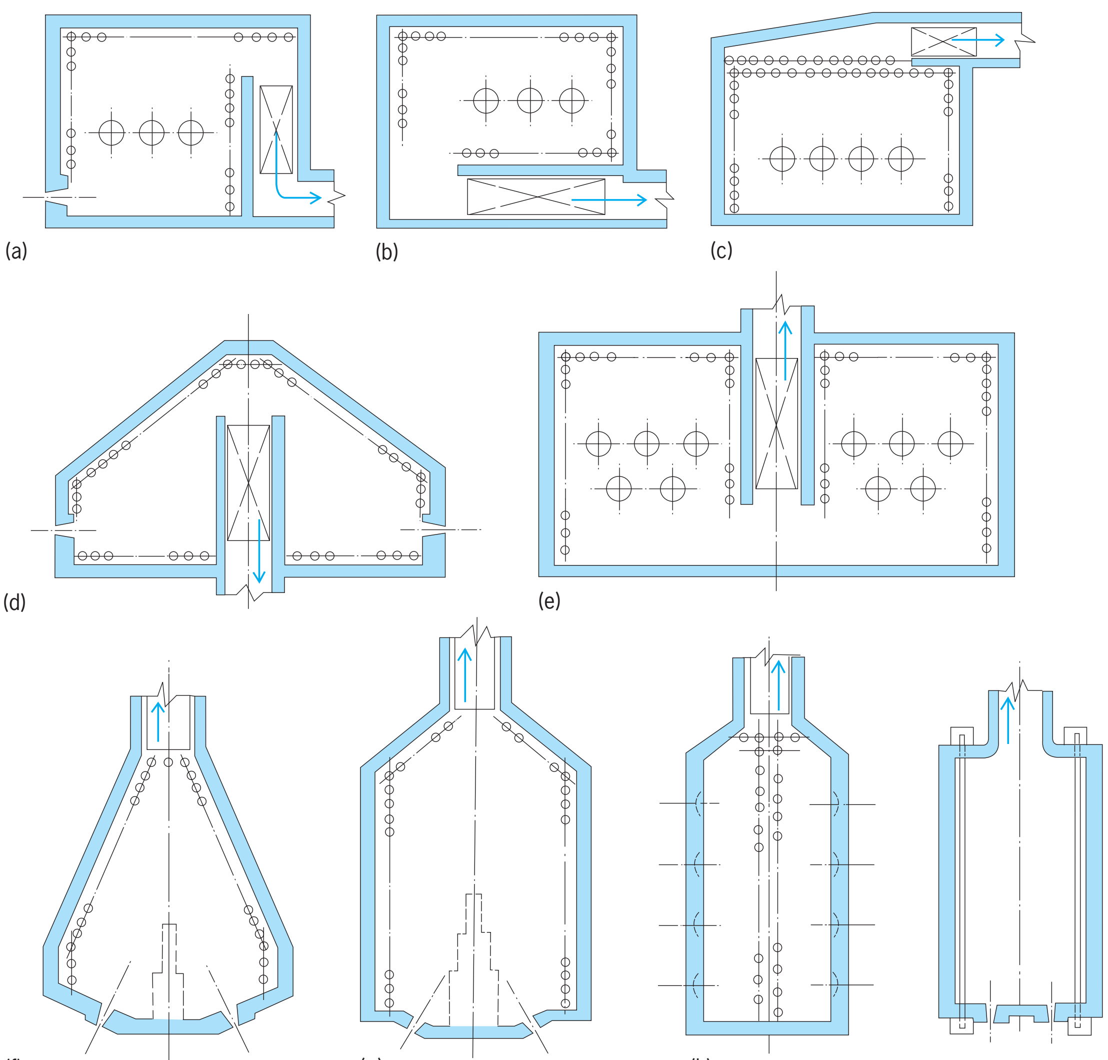


Fig. 1. Typical process furnace designs and characteristic features. (*a*) Radiant, updraft, vertical-convection bank. (*b*) Radiant, downdraft, horizontal-convection bank. (*c*) Radiant, updraft, with high convection effect in radiant roof tubes. (*d*) Large refractory heater, dual radiant sections, sloping roofs. (*e*) Large conventional refinery heater, dual radiant sections. (*f*) Updraft furnace, slanting walls, two parallel coils. (*g*) Updraft furnace, vertical walls, two parallel coils. (*h*) Updraft furnace, tubes in center, burners in wall, and with good heat-intensity distribution. (*i*) Updraft circular, all-radiant furnace, with the burners in the floor.

needed and are given in the literature. *See* COMBUSTION; FOSSIL FUEL.

In order to meet a required thermal efficiency, the heat losses must be appraised. There is a certain loss through the furnace setting which varies with the surface area of the enclosure, the heating capacity, and the atmospheric environment. The wall construction of most types of furnaces, that is, application of insulating material, is such that the specific heat loss, Btu/(ft^2)(h) of outer wall area, is practically the same for all operating temperature levels; hence the heat loss will vary principally only with the furnace capacity. **Table 2** shows the average values that will be encountered in industrial furnace constructions. The setting heat loss ΔH_{sl} can be calculated from the relation $\Delta H_{sl} = \lambda H_{fl}$ where λ is the fraction of heat released, and H_{fl} is the enthalpy of the combustion gas at the theoretical flame temperature.

TABLE 1. Fuel-air mixtures

Type of fuel	Type of burner	Normal range of excess air, %
Natural or refinery gas	Air and fuel gas premixed	5–15
Natural or refinery gas	Air induced by natural draft	10–20
Distillate fuel oils	Steam atomization air by natural draft	15–25
Distillate fuel oils	Mechanical atomization air by natural draft	15–30
Residual fuel oils	Steam atomization air by natural draft	20–35

TABLE 2. Heat loss from setting

Heat liberated, Btu/h*	Loss as a fraction of heat released, λ $\Delta H_{sl} = \lambda H_{fl}$
15,000,000	0.030
30,000,000	0.0275
60,000,000	0.0225
75,000,000+	0.020

*1 Btu = 1.055 kJ.

The flue-gas effluent temperature from the furnace can then be determined in accordance with the postulated thermal efficiency η, as shown in the equation below, where H_{st} is the enthalpy of the combustion gas at the convection-bank exit temperature.

$$H_{st} = H_{fl} \times (1 - \eta) - \Delta H_{st}$$

Heat load. The distribution of heat load between the radiant and convection sections of conventional heaters is associated with the manufacturer's design approach and service requirements. There are all-radiant heaters, and furnaces with tubes arranged mostly for heat absorption by convection. The amount and disposition of heating surface in the combustion or radiant chamber has a distinct influence upon the radiant-heat absorption and the supplementary heat to be recovered by convection for the required thermal efficiency. Therefore, the allocation of heating duties to the radiant and convection sections is more or less an empirical matter. Conventional process heaters are generally designed for approximately 75% thermal efficiency in the United States, where fuel costs are relatively low. Most box-type heaters are designed for moderate radiant-heat-transfer rates; approximately 65–75% of the heat load is carried by the radiant section and 25–35% by the convection bank. In some cases, very high radiant-heat intensities are employed as a matter of design principle, and in these furnaces, the heat absorbed in the radiant section may be only 50% of the total duty.

Radiant section. The present procedure for determining the radiant-heat-transfer rate and heating surface rests on a long series of scientific investigations and evaluation of operating data.

The combustion gases radiate to the solid-body environment, tubes, and refractories. Carbon dioxide and water vapor are the principal constituents with emissive power. Their emissivity depends upon the partial pressure and the thickness of the gas layer, and at 1900°F (1040°C) has a value of approximately 25% of the blackbody radiation intensity. The terms luminous- and nonluminous-flame radiation used in scientific treatises are somewhat misleading; the mere luminosity or light has no significance as far as heat radiation is concerned, because practically all emissivity, and certainly carbon dioxide and water radiation, is in the infrared wavelength region, even at flame temperatures encountered in process heaters. The maximum gas radiation for an infinite value of the product of the partial pressure and thickness of the gas layer, termed black gas, amounts to only 21–33% of the complete blackbody emissivity at temperatures of 2500 and 1100°F (1400 and 590°C), respectively. The mean radiating temperature in the combustion chamber lies within this range.

The consideration of heat transmission by radiation alone, from the combustion gas and refractory walls by reradiation, has not resulted in the derivation of a universally applicable formula for the heat-transfer rate in the radiant section. This is because heat is also transmitted to the tubes by convection, which is on the order of 10% to as high as 35% of the heat absorbed in the radiant section, depending upon (1) the shape of the combustion chamber, (2) the arrangement of tubes, and (3) the type and location of burners. The convection effect is a linear function of the temperature difference (not the difference of temperatures to the fourth power) and can be varied for the same furnace design by the amount of excess air employed. *See* HEAT RADIATION.

Convection heating surface. The rate of heat transfer solely by convection, with flow of gas at a right angle to tubes, increases with the gas velocity and the temperature level, and is greater for small than for large tubes.

There are radiation effects in the convection-tube bank augmenting the heat absorption, especially when the flue-gas temperatures or the temperature differences between tube wall and flue gases are large. These include (1) radiation to the front two rows of tubes, which can be appreciable if they are exposed to the combustion chamber; (2) radiation from the hot gases surrounding the tubes; and (3) the radiation from the refractory enclosure to the tube bank as a whole. *See* CONVECTION (HEAT).

Process furnaces. The design of process furnaces, although following generally the procedure outlined earlier for conventional heaters, requires careful consideration of the transitory state of the fluid being heated.

Vacuum distillation of a temperature-sensitive fluid or the stripping of solvent from the raffinate and extract of a selective solvent-separation process are typical cases in which temperature limitations are encountered in the heating operation. Partial vaporization of the fluid must take place in the tubes under relatively mild heat-transfer rates and tube-wall temperatures, not greatly in excess of the bulk-flowing temperature. Aside from the knowledge and distribution of the combustion-gas side, the proper design of the furnace coil involves exacting pressure-drop calculation for mixed vapor-liquid flow in conjunction with phase-equilibrium determination of the amount and composition of vapor and liquid at any given point in the coil. The pressure drop for mixed vapor-liquid flow is reasonably well established, and vapor-liquid equilibrium relationships are known for many binary and ternary systems. The determination of temperature, pressure, and transition state of the fluid throughout the coil involves stepwise trial-and-error calculations which can best be executed by a modern digital computer for various coil dimensions and heat-intensity patterns. The coil design that satisfies all technical criteria and cost aspects can then be selected from the computer evaluations. *See* DISTILLATION.

Gas reforming is another industrially important heating process involving a chemical change of the fluid. Methane or natural gas and steam are decomposed in the presence of excess steam in tubes filled with a catalyst. The principal products are hydrogen, H_2, carbon monoxide, and carbon dioxide, and maximum conversion of hydrocarbons is attempted. The chemical transformation proceeds at approximately 1400–1450°F (760–790°C) and is highly endothermic, having a heat of reaction of 87,500 Btu/(lb)(mole of methane converted). The tubes are located in the center of the radiant or combustion chamber and are heated from two sides to effect an even circumferential heat-intensity distribution. The average rate of heat transfer is in the order of 18,000 Btu/(ft^2)(h).

Pyrolysis of hydrocarbons, the thermal conversion of ethane, propane, butane, and heavier hydrocarbons to produce olefins and diolefins for the petrochemical industry, is another example of a process furnace. The modern pyrolysis furnace is an outgrowth of the oil-cracking heater, the mainstay of motor-fuel production from petroleum before the advent of catalytic cracking. The design of a pyrolysis furnace must take thermochemical equilibria and reaction rates into consideration in order to realize a high conversion and a product distribution most favorable with respect to yield of olefins. The conversion of higher-molecular-weight hydrocarbons which have a high rate of decomposition involves the problem of attaining the coil outlet temperature dictated by equilibrium relationships. A shorter coil-residence time and correspondingly higher heat-transfer rates are means to this effect, but still the basic principle of moderating the heat intensity in the pyrolysis coil zone, in which the reactant is already in a high state of conversion, must be observed. **Figure 2** gives a typical temperature-pressure and heat-flux pattern for a pyrolysis coil cracking a naphtha boiling in the range 190–375°F (88–191°C) for maximum ethylene yield. The design of such a pyrolysis coil can be executed with greatest precision or refinement by the use of a modern digital computer, permitting the rapid evaluation of all design variables. *See* CRACKING; PYROLYSIS.

Mechanical construction. The design of the furnace setting and application of refractories and insulating materials generally follow standard practice prevailing in boiler-furnace construction. Fully suspended wall construction and the use of lightweight firebrick with bulk densities of 30–65 lb/ft^3 or 480–1040 kg/m^3(as compared with 125–150 lb/ft^3 or 2000–2400 kg/m^3 for fireclay bricks) are generally preferred for process heaters to reduce the heat storage in the setting, start-up, and shutdown time. *See* REFRACTORY.

The burners are carefully selected and must offer great flexibility with respect to firing rate and excess air. The use of a multiplicity of small-capacity burners is rather common for specific process heaters such as gas-reforming and pyrolysis furnaces.

The design of the heating coil for high-temperature service involves the application of alloy steels. Tube materials are usually selected according to the schedule in **Table 3**. Tubes of carbon and low-chrome steel are usually connected by forged-welding return bends, and those of stainless steel by cast 25/12 chrome-nickel alloy steel bends, also weldable to the tubes.

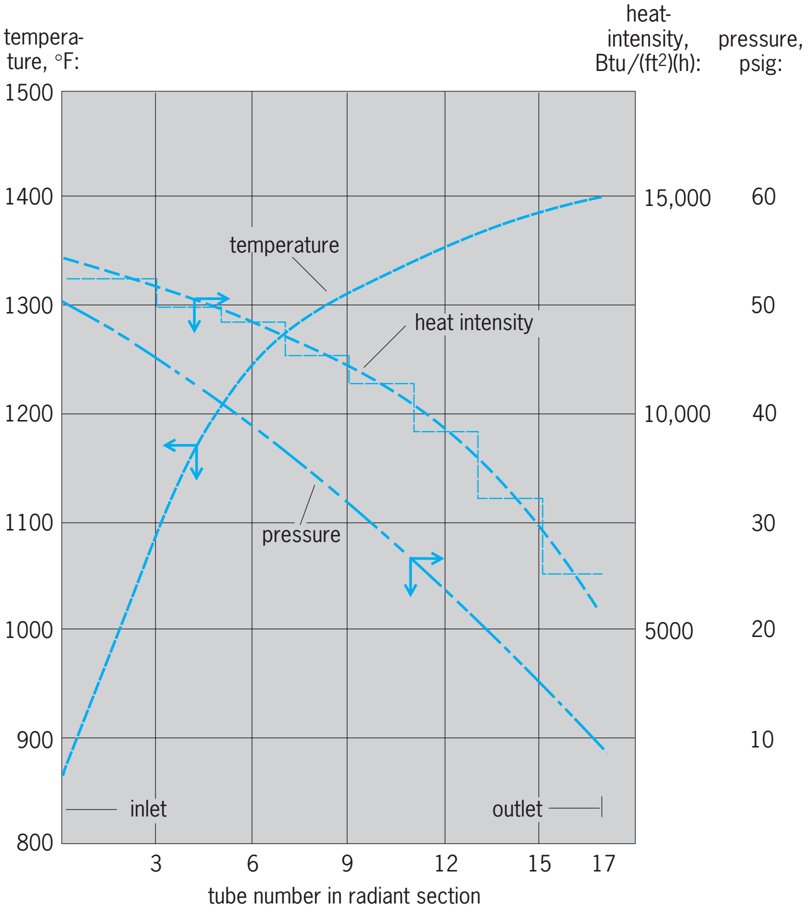


Fig. 2. Typical temperature and heat-flux pattern for naphtha pyrolysis coil. °C = (°F−32)/1.8. 1 Btu/(ft^2)(h) = 3.15 W/m^2. 1 psig = 6895 Pa.

The flue-gas exhaust is normally carried to the atmosphere by a stack of sufficient height to induce a draft equivalent to the pressure drop in the flue-gas passage. For highest heat economy a process furnace can readily be equipped with an air preheater or waste-heat boiler usually requiring air and flue-gas exhaust blowers. Preheated combustion air

TABLE 3. Heating-coil materials

Tube-wall temperature range, °F (°C)	Tube material*
875 (470) and under	Carbon steel, A-161
875–1150 (470–620)	Low-chrome steel, A-200 (1.25–2.25% Cr, 0.5% Mo)
1150–1550 (620–840)	$1/\sqrt{100}$ stainless steel, A-271, type 304
1550–1800 (840–980)	$2/\sqrt{520}$ stainless steel, A-271, type 310

*American Society for Testing and Materials specifications.

results in higher flame and mean effective radiation temperatures which aid in attaining the high radiant-heat-transfer rates required in certain pyrolysis operations. These furnace accessories are provided in the same manner as in boiler design practice. *See* CHIMNEY; GAS FURNACE; HEAT EXCHANGER; OIL FURNACE; PYROMETALLURGY. H. C. Schutt; H. L. Beggs

Bibliography. J. D. Gilchrist, *Fuel, Furnaces, and Refractories*, 1977; D. Q Kern, *Process Heat Transfer*, 1950; R. D. Reed, *Furnace Operations*, 3d ed., 1981.

Fuse (electricity)

An expendable protective device that eliminates overload on an electric circuit. The fuse is connected in series with the circuit being protected. The components of a typical low-voltage high-power fuse are a fuse element or wire, an insulating material support and housing, two metal end fittings, and a filler (**Fig. 1**).

The fuse element is a silver strip or wire that melts when the current is higher than the rated value. The melting of the wire generates an electric arc. The extinction of this arc interrupts the current and protects the circuit. The fuse element is connected to the metal end fittings which serve as terminals. *See* ARC DISCHARGE.

The filler facilitates the arc extinction. The most commonly used filler is sand, which surrounds the fuse element. When the fuse element melts, the heat of the arc melts the sand near the element. This removes energy from the arc, creating a channel filled with the mixture of melted sand and metal. The metal particles from the melting fuse wire are absorbed by the melted sand. This increases the channel resistance, which leads to the gradual reduction of the current and the extinction of the arc. The insulating support and the tubular housing holds the fuse elements and the filler, which also serves as insulator after the fuse has interrupted the current.

The interruption time is the sum of the melting and the arcing time. It is inversely proportional to the current, that is, a higher current melts the wire faster (**Fig. 2**). The fuse operates in a time-current band between maximum interruption time and minimum meeting time. It protects the electric circuit if the fault current is interrupted before the circuit elements are overheated. The arc extinction often generates overvoltages, which produce flashovers and damage. A properly designed fuse operates without overvoltage, which is controlled by the shape of the fuse element and by the filler.

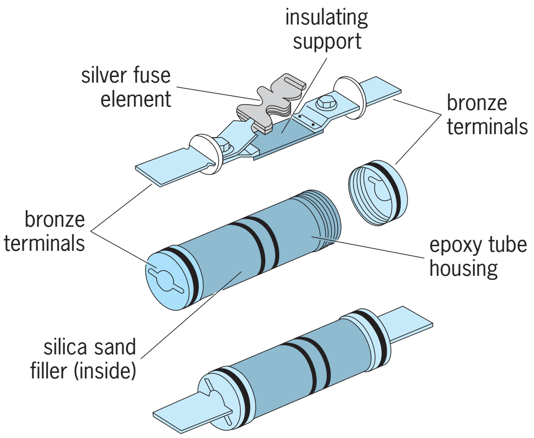


Fig. 1. Power fuse assembly. (*After A. J. Pansini, Electric Power Equipment, Prentice Hall, 1988*)

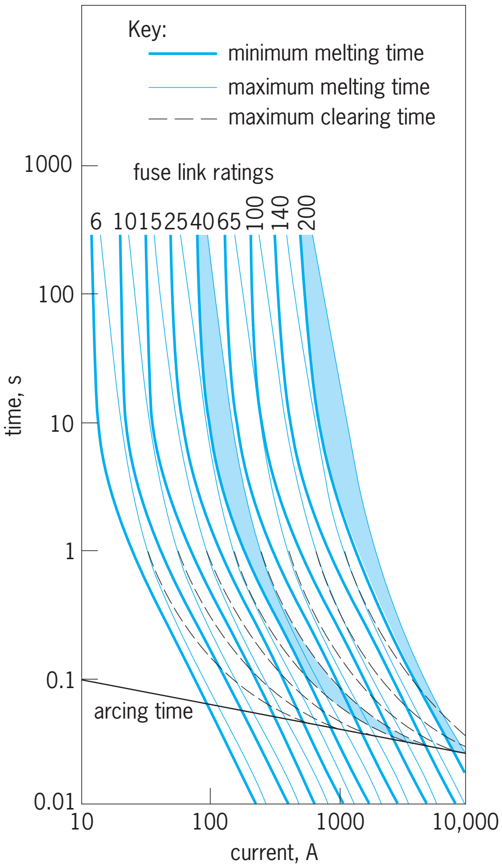


Fig. 2. Minimum and maximum time-current curves for type-K fuse links. (*McGraw Edison Co.*)

Ratings. Fuses protect every kind of electric equipment from the low-voltage direct-current circuits of a motor vehicle to alternating-current circuits of a power distribution system. The fuses are rated to operating voltage and current. The operating voltage is the system voltage; it appears across the fuse after circuit interruption. The voltage rating is between a few volts to 220,000 V. The fuse can carry the rated current continuously. Typical values range from milliamperes to a few thousand amperes. The fusing or melting operation starts around two times the rated current, although the fuse is able to interrupt several thousand amperes of fault current. As an example, the National Electric Manufacturers Association (NEMA) requires a minimum interrupting capacity of 10,000 A from low-voltage ac fuses. Some fuses are able to interrupt 100,000 and even 200,000 A.

Many applications require the connection of fuses in series. A typical example is a low-voltage bus-bar with five outgoing lines each rated 40 A. The bus-bar is protected by a 200-A fuse, and the outgoing lines are protected by 40-A fuses. A line fault should be cleared by the line fuse before the bus-bar fuse starts to melt. This is achieved by coordinating the fuses, which requires that the fuse characteristics are different. Figure 2 shows that the 40-A fuse interrupts any fault current before the 200-A fuse starts to melt. Safety requires at least a 0.4-0.6-s difference in operation time. *See* BUS-BAR.

Classification. Fuses are classified according to whether they carry direct or alternating current and the level of voltage.

Low-voltage direct-current. Typical dc fuses are used to protect circuits in an automobile or in electronic equipment. The modern automobile uses a blade-type fuse that has a ribbon fuse element mounted on a flat substrate.

Electronics equipment uses a cartridge fuse, which consists of a glass tube with two metal caps and a fuse wire or blade between the caps. Both of these fuses are mounted between spring-loaded contacts or held in place by a screwed tubular housing.

In a special category are fuses designed to protect semiconductor devices such as diodes, transistors, and thyristors. These fuses operate extremely fast and interrupt the current before the sensitive semiconductor is damaged.

Low-voltage alternating-current. Typical examples are the ac fuses used to protect domestic electric systems. The fusible element, a silver strip, is mounted in a holder similar to the base of the standard light bulb (**Fig. 3**). Melting starts at the indent or notch on the fusible element. Burned-out fuses can be changed without switching off the supply, because the top of the holder is made of insulating material.

High-power industrial systems are protected by cartridge-type fuses (Fig. 1). The expendable fuse is bolted on two insulated metal blocks or held by spring-loaded contacts.

The starting of an induction motor causes a short overcurrent, three to eight times the motor rated current. Repeated start of the motor should not melt the fuse, but a short circuit in the motor or the blocking of the rotor should melt the fuse immediately. Motor protection requires time-delayed fuses, with longer operation time for moderate overcurrents. This is achieved by increasing the thermal capacity of the fuse element by attaching a blob of solder to the center of an oversized silver wire. The wire delays the melting of the fuse for short durations of overcurrent. Sustained overcurrent melts the solder, which dissolves the silver and initiates fuse operation. Different time-delay fuses are available for motor protection, but none provide full protection, for which a combination of a fuse and a magnetic circuit breaker is required. The fuse provides protection for short circuits and the circuit breaker for overloads.

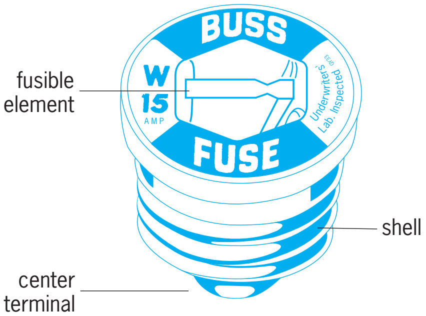


Fig. 3. Common plug-type fuse. (*Bussman Manufacturing Division, McGraw Edison Co.*)

Short-circuit current and a high-power network can be very high. It may produce both mechanical and thermal damage. Special very fast fuses are used for limiting peak current. They interrupt the current before the peak value is reached. Fast operation is achieved by cutting notches or holes in the fusible element. This results in the simultaneous melting of the element at each notch and produces several arcs connected in series. These arcs melt the sand, which removes energy from the circuit faster than the single arc found in the ordinary fuse. The performance of current-limiting fuses can be improved by covering the notches with solder.

The detection of the melted fuse is difficult, but can be facilitated by equipping fuses with mechanical devices that provide visual indication of fuse melting in the form of a colored flag.

High-voltage. High-voltage fuses are extensively used by utilities for protecting distribution systems, and have limited use for transmission systems. The voltage rating is up to 220 kV, and the interrupting capacity may be over 40 kA.

High voltage requires long fuses; and the most frequently used is the expulsion type. The fuse element is installed in an open-ended unfilled tube cartridge. The melting of the element and the arc generates high pressure. The expansion of the high-pressure gases through the open end cools the arc and interrupts the current.

Liquid-filled tubes and short spring-loaded fusible elements are also used. The melting of the fusible element activates the spring, which pulls the terminals apart, extends the arc, and interrupts the current. Some fuses use boric acid, bone fiber, or a liquid solution as a filler. The arc generates water vapor, which helps to cool and quench the arc.

Pole-mounted fuse cutouts are used for the protecting distribution systems. These fuses, mostly of the expulsion type, are mounted in such a way that operation causes the fuse to drop to an open position. This permits the easy detection of the burned-out elements from the ground. **Figure 4** shows a 15-kV fuse cutout, which can be mounted on distribution line poles and used for both line and transformer protection. An additional advantage of the fuse cutout is that in no-load condition the lineman can open the circuit by an insulated stick (hot stick) for repair. The open fuse cutout provides visual proof that the circuit is open.

For the deenergization of a loaded circuit, manufacturers offer a current-interrupting tool with the fuse cutout. The lineman connects the current interrupter in parallel with the fuse. The device interrupts

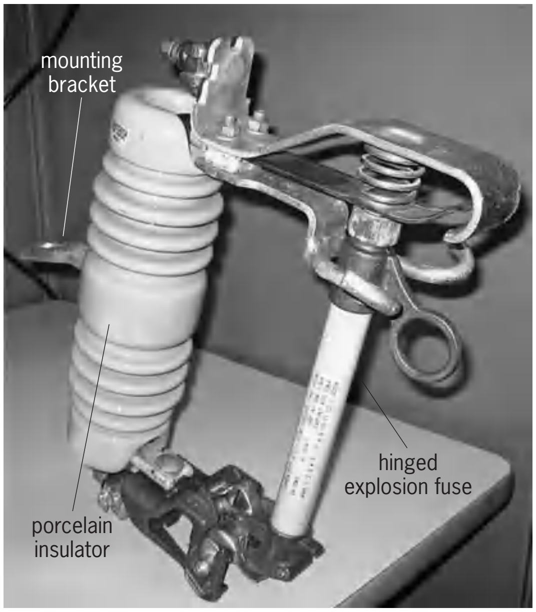


Fig. 4. Typical pole-mounted fuse cutout (15 kV, 100 A).

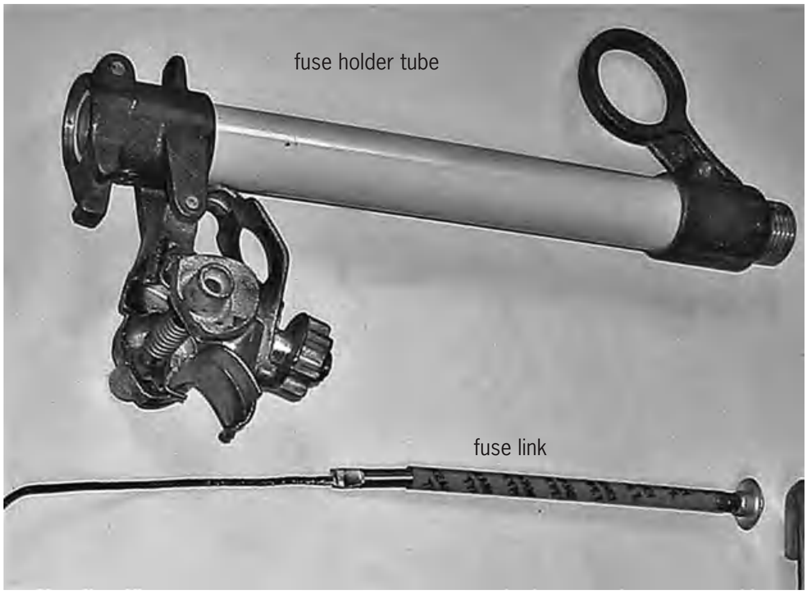


Fig. 5. Fuse holder tube and fuse link.

the load current, and the lineman pulls the hinged fuse to the open position using a hot stick.

The major components of a fuse cutout are mounting bracket, porcelain insulator, and hinged explosion fuse. The fuse contains a fuse holder tube and a fuse link. The fuse link is installed in the plastic fuse holder tube. **Figure 5** shows the fuse holder tube and fuse link.

The fuse link contains one or more fusible elements. **Figure 6** shows the cross section of a typical fuse link. The fusible element is installed in a plastic tube and shunted by a thin wire. A flexible copper wire is connected to the fusible element. The end of the plastic tube is open. The fuse link is installed in the fuse holder. The flexible connecting wire is attached to a spring-loaded lever at the bottom of the fuse holder.

The high short-circuit current melts the fusible element and produces an arc. The arc burns the plastic tube and heats the inner lining of the fuse holder. The high temperature produces high-pressure gases, which exhaust at the bottom of the fuse holder. The high-speed gas jet extinguishes the arc. Simultaneously the spring-loaded connecting wire pulls the burned fusible element out and triggers the fuse-holder opening mechanism. The pivoted fuse holder rotates to the open position.

Overload current melts the thin wire, which produces gases by heating the thin plastic tube surrounding the fuse link. The generated gas extinguishes the arc. The spring-loaded connecting wire pulls the fusible link out and triggers the fuse-holder opening mechanism. The fuse holder swings down to the open position. Most fuses of this type interrupt the current at a zero crossing.

In addition, current-limiting fuses are used for the protection of medium-voltage (15-kV) distribution transformers. These fuses interrupt current safely above three times the rated current but start to melt around one and a half times the rated current. This fuse limits the large short-circuit current because it interrupts the current before the current reaches its peak value.

In order to assure safe current interruption in the full operating range, a type of fuse has been developed with two fuse elements in one housing: a current-limiting element and an expulsion element connected in series. The expulsion element operates at low current, and the current-limiting element operates at high current. Another solution is the use of gas-emitting material in the current-limiting fuse. The long-duration, low-current arc produces gas, which

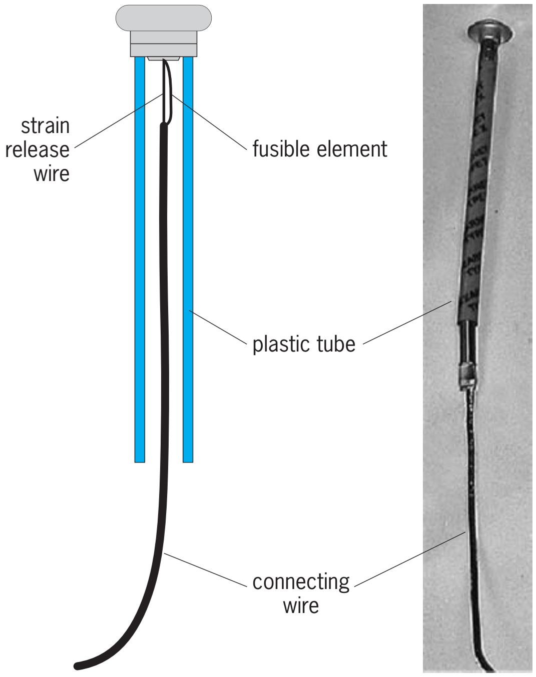


Fig. 6. Cross section and photo of the fuse link.

expands, blows out the arc, and interrupts the current. *See* CIRCUIT BREAKER; ELECTRIC PROTECTIVE DEVICES.
George G. Karady

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; P. Kiameh, *Electrical Equipment Handbook*, McGraw-Hill, 2003; R. E. Loyd, *Electrical Raceways & Other Wiring Methods: Based on the 2005 National Electric Code*, 5th ed., Thomson Delmar Learning, 2005; A. J. Pansini, *Electric Power Equipment*, 1988; B. Peoples, AC versus DC fusing, *Electrical Insulation Conference and Electrical Manufacturing & Coil Winding Conference: 2001 Proceedings*, pp. 239–243, Oct. 16–18, 2001; R. Peters, *Electrical Basics*, Sterling Publishing, 2003; A. Wright and P. G. Newberry, *Electric Fuses*, 3d ed., 2004.

Fused-salt phase equilibria

Conditions in which two or more phases of fused-salt mixtures can coexist in thermodynamic equilibrium. Phase diagrams of these equilibrium conditions summarize basic knowledge about fused salts. Numerous advances in the technologies which are based on high-temperature chemistry have become possible through the increase in knowledge about fused salts. The increasingly significant role of fused salts in industrial processes is evident in the widening application of these materials as heat-transfer media, in extractive metallurgy, in nonaqueous reprocessing of nuclear reactor fuels, and in the development of nuclear reactors which create more fuel than they consume (breeder reactors); moreover, these technologies are all direct outgrowths of research and development with fused salts. *See* NUCLEAR REACTOR; PHASE EQUILIBRIUM.

Fused-salt media were of great interest before 1900 for the electrolytic reduction of cations to metals. The bulk of research on fused-salt phase equilibria has stemmed from interest in electrometallurgical processes from that time until the mid-twentieth century. In the 1950s and 1960s the U.S. Atomic Energy Commission's programs showed that molten-salt mixtures may serve successfully as nuclear reactor fuels, fuel solvents, coolant fluids, moderators, nuclear reactor breeder blankets for the preparation of plutonium or uranium-233, high-temperature bearing lubricants, and fuel-reprocessing media.

Fused-salt mixtures find application in technology when the need arises for liquids which are stable at high temperatures. For most applications, suitably low melting temperatures and low vapor pressures are primary considerations. To some extent these requirements are conflicting, because salts which are useful in obtaining low freezing temperatures often tend to have appreciable covalent character and therefore to exhibit unfavorably high vapor pressures.

As a special class of liquids, one which is composed entirely of positively and negatively charged ions undiluted by weak-electrolyte supporting media, fused salts are used in many different types of research. For example, advances in solution theory, thermodynamics, and crystal chemistry have come about through studies of fused-salt systems. *See* FUSED-SALT SOLUTION.

A close connection between fused-salt phase diagrams and geochemistry exists and stems from the model principle developed by V. M. Goldschmidt, who noted that isomorphic structures are assumed by ions of the same proportionate size and stoichiometric relations but of different charge. Thus the fluorides of beryllium, calcium, and magnesium, for example, are structural models for silicon dioxide (SiO_2), titanium dioxide (TiO_2), and zirconium dioxide (ZrO_2). The fluoride structures are referred to as weakened models because of the smaller electrostatic forces resulting from smaller ionic charges; they have been useful for comparisons with oxide and silicate systems. According to Goldschmidt's interpretation, saltlike materials were derived from components such as water (H_2O), carbon dioxide (CO_2), sulfur trioxide (SO_3), chlorine (Cl_2), and fluorine (F_2), which were volatilized from molten magmas as they crystallized. Crystallization equilibria in fused-salt systems therefore provide a convenient way to study the mechanisms occurring in the formation of igneous rocks. Such investigations were once pursued vigorously, producing the basic techniques which have been employed routinely in phase studies, essentially since the enunciation of the phase rule by J. W. Gibbs. *See* IGNEOUS ROCKS; LITHOSPHERE; PHASE RULE.

Methods of investigation. For many years scientists in the United States and elsewhere expended a great deal of effort in attempts to predict the phase relationships of uninvestigated systems; however, the necessary thermochemical data which were required, such as heat capacities of the constituents, the variation of this property with temperature, the heats of fusion, and the free energies of mixing, were not available. It soon became evident that the acquisition of these data was at least as arduous a task as the determination of the actual equilibrium behavior, and that quantitatively the phase relationships predicted from the thermochemical data were too imprecise to make the method attractive. As a result, it has not yet become routine to estimate with practical accuracy the behavior of untested systems. Accordingly, the determination of equilibrium phase diagrams has remained principally in the region of experimental science. To be sure, numerous correlations are available which can assist the experimentalist in minimizing tasks such as determining the interrelationships of crystal structure, thermodynamic quantities, and the predictable degree of ideality of some ionic systems. Such information makes it possible for phase studies to be conducted with greater efficiency and rapidity, and therefore tends to maintain the field as experimental. The largely experimental character of fused-salt phase equilibria is even more a consequence of the development of experimental techniques which permit the delineation of phase relationships in fused-salt systems far more accurately and rapidly than was previously possible.

Innovations in the original means of study have brought about the application of a wide variety of experimental techniques to studies of heterogeneous phase equilibria. These include differential thermal analysis, visual methods for direct observations of melting and freezing behavior at high temperatures, filtration of partially molten mixtures, high-temperature centrifugation, hot-stage microscopy, and high-temperature x-ray diffraction. Electromotive-force measurements and vapor-pressure determinations have also been used with success in the study of heterogeneous equilibria in salt systems.

One of the most important developments among the experimental devices used for solid-state phase studies was the automatic x-ray diffractometer for single crystal structure studies. With automatic diffraction equipment, the average period for determination of new structures is about one-tenth that required previously. With the assistance of computerized equipment, it is commonplace to measure automatically the intensities associated with several thousand reflections from crystal planes and to deduce the structures with significantly greater accuracy than was obtained previously. This has a special function in phase studies: By concurrently determining the equilibrium relations in a system and the structure of the crystalline solids which are found in the system, the stoichiometry, compositional variability, coordination chemistry, and so on, are established unequivocally.

Types of equilibria. In use, fused salts are generally employed under conditions where temperature and composition are varied extensively. Drastic changes in equilibrium phase relationships are thus commonplace. In order to understand how these salts behave chemically, it is necessary to gain a rudimentary knowledge of heterogeneous equilibria. Pragmatically, this approach need not entail a complete nor even extensive study of the systems involved; it will be useful as a framework which prescribes the conditions that can be employed in use and in related investigations with fused salts. Characterization of salt phase equilibria is simple and straightforward with binary and ternary systems, but may become increasingly intricate and of decreasing utility as the effect of additional components must be considered. Considerations here are therefore limited to condensed equilibria in binary and ternary systems.

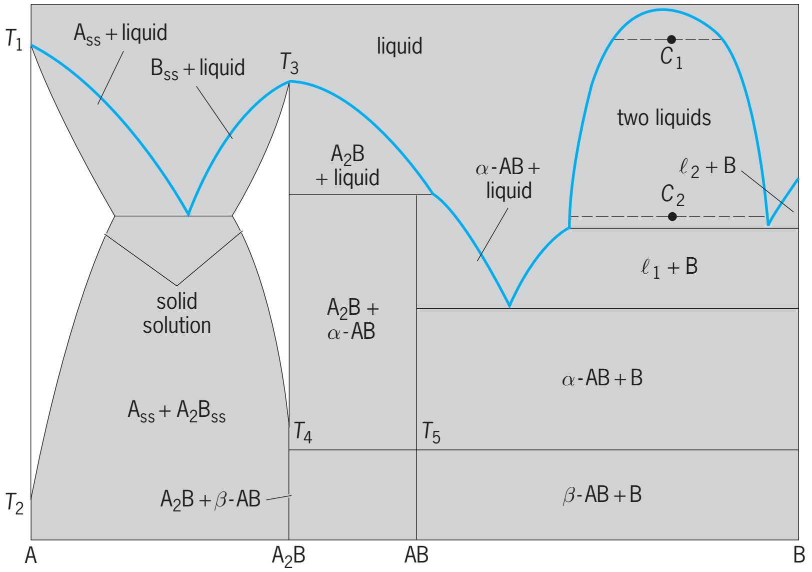


Fig. 1. Diagram of typical phase equilibrium relationships in condensed binary systems of molten salts. Here ℓ is liquid and ss is solid solution.

Binary systems. Consider first the binary system A-B. As drawn (**Fig. 1**) the phase diagram indicates that the components A and B react to form the intermediate compounds A_2B and AB. Limited miscibility in the solid state occurs when mixtures of A and A_2B are present at equilibrium at temperatures between T_1 and T_2, and T_3 and T_4. The compound AB melts incongruently to A_2B and liquid. Crystals of AB are dimorphic; at temperatures above T_5 they exist in a crystalline state which transforms, on cooling below temperature T_5, to a less symmetric (lower-energy form) crystalline state. Liquids formed from mixtures of AB and B are immiscible at compositions such as C_1 and C_2, which separate into immiscible liquids of composition described by horizontal tie lines in the two-liquid region shown in Fig. 1.

Ternary systems. The temperature-composition relations in condensed ternary systems are generally represented by a polythermal projection of the liquidus surfaces. This is a projection parallel to the temperature axis on the triangular composition plane; it shows, therefore, the various parts of the liquidus surface or surfaces as viewed from the direction of high temperature. Such diagrams indicate the compositions of liquids saturated with one or more solids. At the invariant reaction points, liquids are in equilibrium with three solids. The direction of temperature change is shown by arrows on boundary curves, and liquidus temperatures are shown by isothermal contours.

The hypothetical ternary system shown in **Fig. 2** displays behavior which is typically described by three-component phase diagrams. Assume that invariant equilibrium points in the phase diagram A-B-C will ultimately be located as shown in Fig. 2 and that there are no intermediate compounds formed from components B and C; a single intermediate compound AC_2, melting congruently, occurs in the system A-C; the congruently melting compound AB_2 and the incongruently melting compound A_2B are formed from components A and B. A careful examination of several crystalline mixtures cooled from the liquid state can be expected to furnish much information about equilibria in the system A-B-C. Some examples are given below.

1. All crystallized specimens having compositions within the triangle AB_2-B-C contain only the pure solid phases AB_2, B, and C. These results indicate that the compounds AB_2, B, and C form a subsystem, that AB_2 probably melts congruently, and that the section

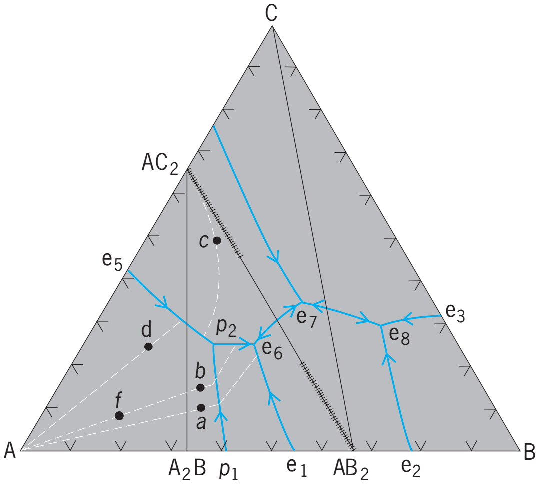


Fig. 2. Typical phase equilibrium relationships in condensed ternary systems of molten salts.

AB_2-C is a quasi-binary system. Under these circumstances, a single invariant point, the eutectic e_8, can occur in AB_2-B-C. It may be possible to estimate its approximate composition by examination of the crystallized mixtures. Morphological data, obtained from petrographic or metallographic methods, may reveal the domains of the primary phases AB_2, B, and C, and the composition of the eutectic.

2. Assume that AC_2 and AB_2 form extensive but limited solid solutions exhibiting maximum solubility at the solidus, as shown in Fig. 2. In general, lattice constants of AB_2 and AC_2 will change as a function of solute concentration. Therefore, x-ray data obtained from crystallized specimens having compositions intermediate between AB_2 and AC_2 can be examined for lattice-spacing changes and correlated employing the parametric method. Some exsolution (precipitation of crystals from the solid solution) is to be expected as temperatures are lowered. The exsolution curve often exhibits considerable slope and precludes the possibility of obtaining data which indicate concentrations of maximum solubility of AB_2 and AC_2 except by static methods.

3. Occurrence of solid solutions of AB_2 and AC_2 indicates that inferences regarding the subsystem AB_2-AC_2-C can be made from phase analyses of mixtures having compositions in triangle AB_2-B-C in addition to those in AB_2-AC_2-C. Some mixtures in triangle AB_2-AC_2-C which crystallize primary-phase AC_2 solid solutions will solidify before the liquid fraction attains the composition e_7 and may not contain detectable concentrations of AB_2 or C. Mixtures which crystallize primary-phase AB_2 solid solutions will behave similarly, producing possibly undetectable amounts of AC_2 or C. The identity of subsystem AB_2-AC_2-C may thus be equivocal without the conclusion based on other data that AB_2-C is a quasibinary.

4. Much information about phase behavior with the triangles A_2B-AC_2-AB_2 and A-A_2B-AC_2 can be obtained from x-ray analyses of specimens crystallizing within triangle A_2B-AC_2-AB_2. The solid-solution formation along AB_2-AC_2 and the melting relations of A_2B will cause mixtures with A_2B-AB_2-AC to undergo a variety of crystallization reactions. Depending on the character and extent of the solid solution and the melting-freezing behavior of A_2B, x-ray analysis may furnish very informative phase data. Crystallization in polycomponent systems often results in the formation of final products which coexist in an unstable equilibrium. Identification of the crystalline phases in the final crystallized mixture provides useful inferences as to the nature of the equilibrium phase diagram. *See* EUTECTICS.

Applications in research and development. Two general categories encompass the principal activities in phase equilibrium investigations, those which center about technological applications and those which are essentially fundamental science.

The principal applications in which fused salts are employed in technology and research are listed below. As is implied in the foregoing paragraph, the division between technology and research for a number of the listed categories is artificial.

Technology

1. Reductive extraction processes for metals production
    a. Electrochemical processes: (1) Sacrificial anode methods: tantalum (Ta), niobium (Nb), tungsten (W), and molybdenum (Mo). (2) Electrowinning methods: lanthanides, aluminum (Al), transition metals.
    b. Pyrochemical processes: lanthanides actinides
2. Working fluids for molten-salt breeder reactor development
    a. Fuels, based on uranium fluoride (UF_4), thorium fluoride (ThF_4), plutonium fluoride (PuF_3)
    b. Blankets for breeder reactors
    c. Coolant salts
3. Solvents for use in volatility processes for recovery of uranium from spent fuel elements
4. Nonnuclear heat-exchange media: nitrates, fluoroborates, carbonates
5. Application to regenerative fuel cells
6. Metal alloys–salt compatibility
7. Development of analytical chemical procedures

Research

1. Coordination chemistry
    a. Lanthanide-actinide model relationships
    b. Complex ions in solutions
    c. Electrochemical research
    d. Transition-metal chemistry
    e. Structure studies
2. Thermodynamics of solution behavior
    a. Freezing-point depression
    b. Liquid-liquid immiscibility
    c. Reciprocal systems
3. Optically active crystals
    a. Lasers
    b. Electronic theory

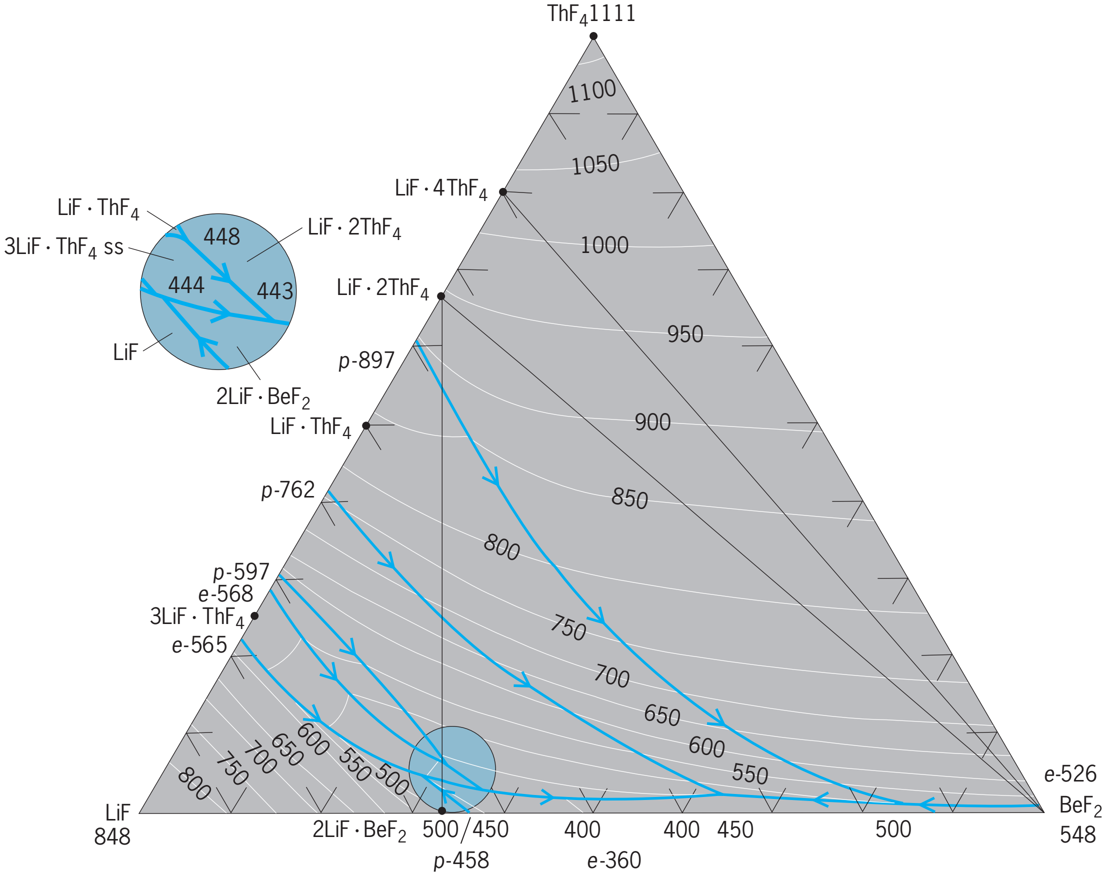


Fig. 3. The system LiF-BeF$_2$-ThF$_4$. Temperature is in °C. °F = (°C × 1.8) + 32. (*After R. E. Thoma et al., Phase equilibria in the systems BeF$_2$-ThF$_4$ and LiF-BeF$_2$-ThF$_4$, J. Phys. Chem., 64:865–870, 1960*)

4. Crystal growth
   a. Single crystals of oxides and fluorides for structure studies
   b. Zone refining
   c. Phonon scattering
5. Metal-metal halide systems
6. Source of new types of molten-salt research

Perhaps the principal factor which distinguishes technological and research applications of molten-salt phase equilibrium studies is that the materials problems in technology often necessarily involve multicomponent systems, whereas the number of components in fundamental research studies is limited by choice. In technology the processes are frequently highly complex reactions involving simultaneously metal, gas, salt, and oxide or other phase reactions as well as oxidation-reduction equilibria. One application, involving a variety of fused-salt systems, has been in molten-salt breeder reactor technology.

In the United States during the 1960s, efforts to develop advanced-design nuclear reactor technology included those devoted to a single-region fluid-fuel reactor. Advances in the chemistry of processing molten salts to remove protactinium and the lanthanides indicate that development of a large, 1000-MW (electrical), single-fluid molten-salt breeder reactor would be feasible. Thermal energy generated in the reactor fluid would be transferred to a secondary coolant-salt circuit, which couples the reactor to a supercritical steam cycle. The results of design studies showed that the salt system should be composed of lithium fluoride (^7LiF), beryllium fluoride (BeF_2), ThF_4, and $^{235}UF_4$, with concentrations optimized by nuclear considerations and constraints imposed by phase equilibria and physical properties. The phase diagram of the ^7LiF-BeF_2-ThF_4 system (**Fig. 3**) then dictated the choices which were available for selecting the optimal composition of the salt mixture. For minimum liquidus temperatures, minimum viscosity of the molten salt, and optimal nuclear performance, the composition ^7LiF-BeF_2-ThF_4-$^{233}UF_4$ (68:20:11.2:0.2 mole %) was selected provisionally. Although development of the technology was terminated, its feasibility remains viable. Any future development of the fuel and reprocessing technology will entail optimizations that consider the phase behavior of fused-salt systems that are candidates for application. *See* NUCLEAR POWER. R. E. Thoma

Bibliography. J. Braunstein et al., *Advances in Molten Salt Chemistry*, vols. 1–4, 1971–1981; L. Eyring (ed.), *Progress in the Science and Technology of the Rare Earths*, vol. 2, 1966; D. Inman and D. G. Lovering (eds.), *Ionic Liquids*, 1981; G. J. Janz, *Molten Salts Handbook*, 1967; I. M. Kolthoff (ed.), *Treatise on Analytical Chemistry*, pt. 1: *Theory and Instrumentation*, 1961; J. Lumsden, *Thermodynamics of Molten Salt Mixtures*, 1966; G. Mamantov et al. (eds.), *Advances in Molten Salt Chemistry*, vols. 5 and 6, 1983–1987; G. Mamantov and R. Marrasi (eds.), *Molten Salt Chemistry: An Introduction and Selected Applications*. 1987.

Fused-salt solution

A nonaqueous solvent system particularly useful in the production of reactive metals and in the study of the coordination chemistry of the elements. Fused salts are a large class of liquids that are composed largely of ions. Many simple inorganic salts melt at high temperatures (greater than 600°C or 1100°F), forming liquids which have high specific ionic conductivities (1–6 $\text{ohm}^{-1}\ \text{cm}^{-1}$). There are, however, a number of exceptions; for example, the electrical conductivity of $AlCl_3$ decreases sharply upon melting due to the formation of the molecular liquid Al_2Cl_6. Mixing molecular fused salts with truly ionic salts results in the formation of so-called potential electrolytes. Each molten salt system constitutes its own solvent system, and needs the establishment of its own electromotive force series and requires, in some cases, the development and use of special reference electrodes. *See* ELECTROLYTIC CONDUCTANCE; FUSED-SALT PHASE EQUILIBRIA; HIGH-TEMPERATURE CHEMISTRY; SOLVENT.

The use of binary or ternary melt compositions results in liquids that typically have much lower melting temperatures than do the starting constituents and somewhat lower specific ionic conductivities than do pure salts. The choice of melts for use as solvents frequently is based on considerations such as the availability and cost, the lowest melting temperature attainable, and the ease of purification of the solvent, as well as the width spanning the electrochemical stability and the spectroscopic transparency of the melt. Several molten-salt solvents, such as the LiCl-KCl eutectic and the equimolar $NaNO_3$-KNO_3 melt, have been extensively studied, and many of their physical and chemical properties are well known. Fused salts used in the industrial production of reactive metals are usually multicomponent systems, with each component added to the melt in order to improve a specific property. Many melt systems, particularly ternary and compositions that are more complex, have been only partially characterized; their physical properties are estimated via modeling by using the data available for less complex systems, such as the binary component melts.

Applications. Fused-salt solutions have been used extensively as electrolytes in the production and refining of aluminum (oxyfluoride solutions), magnesium (chloride solutions), lithium (chlorides), and calcium (chlorides) metals. Fused salts have also been used as solvents in the production by thermite (reduction) reaction of high-melting metals such as titanium, zirconium, niobium, tantalum, rare earths, actinides, and alkali metals. Some useful high-temperature new materials, such as titanium diboride and lanthanum hexaboride, also have been synthesized in high purity using fused-salt solutions. One of the largest commercial applications of fused salts is for the heat treatment of metals and alloys, where fused salts are used as baths for their heat dissipation capacity.

Emerging applications. Fused salts have been investigated for use in rechargeable batteries and in fuel cells because they allow the use of higher currents than aqueous systems.

Potential applications. Fused-salt systems are well suited for the environmentally sound treatment of wastes and recycling processes. Fused salts also have been investigated for the formation of synthetic carbonate rocks as a potential permanent storage form of the greenhouse gas carbon dioxide. In addition, low-temperature fused-salt systems, called room-temperature molten salts, can provide the advantages of both the fused salts and the aqueous systems. Room-temperature fused salts consist of a mixture of metal halide salts and organic halides. The most thoroughly studied systems contain aluminum trichloride and alkylpyridinium halides or alkylimidazolium chlorides. Room-temperature molten-salt systems are suitable for electroplating reactive metals.

Chloroaluminate systems. For some melt systems, such as those containing aluminum halides, the melt composition can be adjusted to change the acid-base properties of the solvent, which determine the chemistry of the dissolved solutes. The $AlCl_3$-NaCl system has been studied extensively and will be used to illustrate the concepts involved. An acid (Lewis acid) in this system is defined as a chloride ion acceptor and $pCl^- \equiv -\log [Cl^-]$. (This concept is similar to pH.) The modified Lewis acidity is frequently used in discussing these systems. Raman spectral measurements have shown that the $AlCl_4^-$ ion is the predominant anion in the equimolar $AlCl_3$-NaCl melt. As the $AlCl_3$/NaCl molar ratio is made greater than one, the melt becomes acidic and the concentration of the $Al_2Cl_7^-$ ion increases. Near the equimolar point, the dominant equilibrium is that shown in reaction (1).

$$2AlCl_4^- \rightleftharpoons Al_2Cl_7^- + Cl^- \qquad (1)$$

The mole fraction equilibrium constant for this equilibrium has been determined from potentiometric measurements as 1.1×10^{-7} at 175°C (347°F). The most basic sodium chloroaluminate system corresponds to a NaCl-saturated melt in which the pCl^- is constant in the presence of solid NaCl and is equal to 1.1 at 175°C (347°F). Two other equilibria [reactions (2) and (3)] become important as the melt acidity is increased by increasing the $AlCl_3$ content. Reaction (1) is analogous to the dissociation of water [reaction (4)]. Similar equilibria have been proposed for other

$$3Al_2Cl_7^- \rightleftharpoons 2Al_3Cl_{10}^- + Cl^- \qquad (2)$$

$$2Al_3Cl_{10}^- \rightleftharpoons 3Al_2Cl_6 + 2Cl^- \qquad (3)$$

$$2H_2O \rightleftharpoons H_3O^+ + OH^- \qquad (4)$$

melt systems containing a covalent halide, such as $FeCl_3$ or $SbCl_3$.

A major difference between water (and related solvents) and the chloroaluminate solvent system is the ability of the latter to stabilize unusually low oxidation states of metals in acidic melts. Thus ions such as Cd_2^{2+}, Bi^+, and Hg_3^{2+} are quite stable in $AlCl_3$-rich melts. This unusual chemistry has been attributed to the small tendency for disproportionation of the low oxidation state. Such disproportionation is favored by large concentrations of halide ions which complex preferentially the higher oxidation state. For example, reaction (5) is favored by the presence of chloride ions; the chloride concentration can be made small by adding a Lewis acid such as $AlCl_3$. The chemistry in molten chloroaluminates bears considerable similarity to superacid systems, such as HSO_3F-SbF_5. Thus sulfur, selenium, and tellerium in both media form homopolyatomic cations such as S_8^{2+} and Se_4^{2+}, as well as poorly characterized radical species. *See* CHEMICAL EQUILIBRIUM; SUPERACID.

$$Cd_2^{2+} \rightleftharpoons Cd + Cd^{2+} \tag{5}$$

It should be stressed that small or moderate changes in melt composition can have a drastic effect on the chemistry of solutes in the chloroaluminate melts. For example, Ti(II) is a stable oxidation state in melts quite rich in $AlCl_3$, such as $AlCl_3$-NaCl (65–35 mole %). In less acidic melts, Ti(II) forms a precipitate, while no evidence for the formation of divalent titanium has been obtained in melts that contain an excess of NaCl. Similarly in alkali fluoride melts, only Ti(IV) and Ti(III) exist; the reduction of trivalent titanium leads directly to titanium metal. The use of fluoride media results in greater stability of the high oxidation states because many metal ions in these oxidation states form very stable complexes with fluoride ions.

It is interesting to contrast the reduction of refractory metal ions, such as those of niobium, tantalum, and tungsten, in the basic alkali fluoride media—for example, the LiF-NaF-KF eutectic (Flinak)—and in the acidic chloroaluminate melts. In Flinak, the reduction proceeds through several steps (some of which are irreversible) to the metal, which may be obtained in coherent nondendritic form. In $AlCl_3$-rich chloroaluminate melts, metal clusters, such as W_6Cl_{12}, are readily formed and the formation or plating of the metal may not be possible. On the other hand, the cluster species that is formed may be of interest in other applications, such as catalysis. It has been reported that $Ir_4(CO)_{12}$ in acidic $AlCl_3$-NaCl melts catalyzes the methanation reaction (formation of CH_4 from CO and H_2).

Catalysts. One use of fused salts is as media for organic reactions. Not only does the fused-salt environment provide for better thermal control of the reaction (heat dissipation is readily possible), but it also may serve as a catalyst. For example, molten $SbCl_3$ and $ZnCl_2$ have been found to be effective hydrocracking catalysts for coal. It has been found that polycyclic hydrocarbons, such as anthracene, undergo several types of reactions in $SbCl_3$ melts, including formation of radical cations and protonated species that react further to form condensed systems, such as anthra[2.1-*a*]aceanthrylene. Other examples are the chlorination of benzene in NaCl-ACl_3-$FeCl_3$, the formation of methyl chloride in KCl-$ZnCl_2$, the synthesis of ethyl chloride in $AlCl_3$-KCl-NaCl, and the preparation of silicon hydride in KCl-LiCl. *See* CATALYSIS.

Batteries and fuel cells. Another technological area of great interest in which molten salts play a key role is that of advanced batteries and fuel cells. LiCl-KCl eutectic is the solvent in the rechargeable Li(Al)-FeS (or FeS_2) battery, and sodium polysulfide melts are used in the sodium/sulfur battery which operates at about 350°C (660°F). The rechargeable cell Na/Na^+ conductor/S in $AlCl_3$-NaCl has an open circuit voltage of 4.2 V at 200–250°C (390–480°F) as well as high energy densities. Molten carbonates, such as the ternary Li_2CO_3-Na_2CO_3-K_2CO_3 melt, are used in fuel cells that use H_2/CO mixtures and oxygen as electrochemical fuels. *See* BATTERY; FUEL CELL.

Georges J. Kipouros; Gleb Mamantov

Bibliography. R. J. Gale and D. G. Lovering (eds.), *Molten Salts Techniques*, vols. 1–4, 1991; D. G. Lovering (ed.), *Molten Salt Technology*, 1982; J. Plambeck, *Encyclopedia of Electrochemistry of Elements X: Fused Salt Systems*, 1976; P. Wasserscheid and T. Welton (eds.), *Ionic Liquids in Synthesis*, 2003.

Fuselage

The component of an aircraft that provides the payload containment and the structural connection for the wing and the empennage (tail assembly). The fuselage and the wing are major structural components of an aircraft. The fuselage is the counting structure for the horizontal and tail surfaces that provides stability as well as the means of introducing pitch and yaw control to the aircraft. For some aircraft like fighter and private aircraft, the fuselage houses the engine or engines. The nose or tail gear and the main landing gear are often attached to the fuselage structure. *See* WING.

The fuselage is the main payload structure for all aircraft. The payload can consist of only crew of pilot, electronics, and armament for a fighter, or can contain crew, passenger attendants, passengers, cargo, and baggage for a commercial passenger aircraft. All high-performance jet aircraft obtain optimum performance at high altitudes. This requires a pressurized fuselage, whether it is only a pilot compartment or the total fuselage of a commercial passenger aircraft. The pressurization requirements, the payload support, the wing, empennage, and landing gear connections, and the flight regime establish the structural loading environment for the design of the structural elements of the fuselage. For each aircraft configuration or type, the aircraft's functional requirements will dictate the location, shape, and efficiency of these structural elements.

Configuration. The fuselages of fighter aircraft are usually designed to provide a minimum envelope for the pilot, engines, fuel, electronics, and armament, with the additional requirement of easy access for combat maintenance. By maintaining a minimum external envelope, the fuselage structural weight and aerodynamic drag are minimized. A similar minimum envelope is usually maintained for bomber aircraft, with additional internal volume being required for bomb payload. In both these aircraft the exteriors of the fuselage are tailored carefully to maximize high-speed aerodynamic performance. This fuselage aerodynamic tailoring results in complex external contours and internal structure which is usually more expensive to produce than the structure found in cargo, commerical, or private aircraft. *See* MILITARY AIRCRAFT.

Cargo aircraft fuselages are designed to maximize internal volume and to provide maximum convenience in cargo loading. For military cargo aircraft, the fuselage configurations are basically all similar to those shown in **Fig. 1**.

The fuselage configurations for all large commercial passenger aircraft are established by passenger payload requirements. For those cargo aircraft which are derivatives of commercial passenger aircraft, the fuselage configurations are also controlled by the passenger requirements. The passenger aircraft fuselage configurations are generally divided into two separate sections: the passenger compartment above the floor, and a lower belly cargo compartment. The design objective for private aircraft fuselages which are not pressurized is to provide a simple envelope which is easy to manufacture at minimum cost.

Material and fabrication. The history of the construction of aircraft fuselages has evolved through the early wood truss structural arrangements to the current metal semimonocoque shell structures. Currently a majority of aircraft fuselages are fabricated from aluminum alloys and are produced by a process of automatic machining of the skins and stringers, with much of the assembly being done by automatic drilling, countersinking, and fastener installation. In some areas, adhesive bonding is used as a means of attaching doublers to reinforce skin panels. The Fokker F-27 and F-28 commercial aircraft use extensive adhesive bonding in fuselage construction. Adhesive bonding is also used in the fabrication of the honeycomb sandwich structure, which has been used in the fuselage of General Dynamic's F-111 fighter.

In many of the high-performance aircraft, such as fighters and bombers, extensive use is made of titanium and high-strength steel. The use of numerical machining has allowed the application of these materials to complex bulkheads, fittings, major structural elements, and attachment points. In the fabrication of titanium, there are methods which use diffusion bonding and superplastic forming to reduce the cost

Fig. 1. Military cargo aircraft, showing fuselage configurations. (*a*) C-130 Hercules, (*b*) C-141 Star Lifter. (c) C-5A Galaxy. (*Lockheed Corp.*)

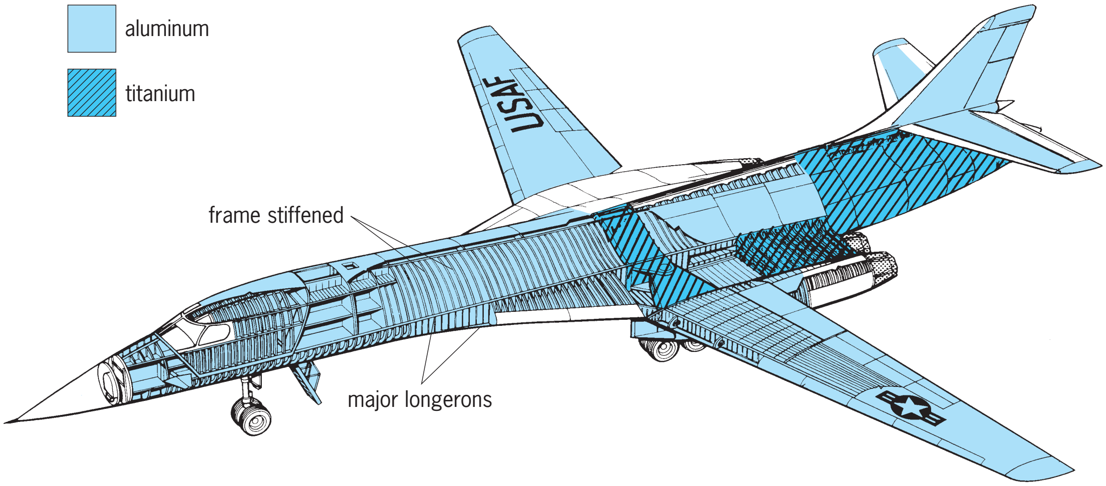


Fig. 2. B-1 bomber fuselage structural configuration, based on frame-stiffened skins with major longerons.

of titanium part fabrication. Other manufacturing techniques include the use of castings, extrusions, forgings, and welding.

Structural configuration. The fuselage structural configuration of modern aircraft is dominated by two generic types of shell structures. One of these consists of a frame-stiffened skin construction with major longerons supporting the bending loads. The other structural configuration is composed of stringer-stiffened skins supported by frames. These two configurations are utilized separately or in combination to produce the most efficient structural configuration.

Fighters and bombers are usually configurated of complex aerodynamic shapes which must provide maximum access to equipment and armament. The best structural design of these fuselages, for functional purposes, is the frame-supported skins with a few major longerons. For this structure configuration the skins carry shear, while the major bending material is located in either the major longerons or the keel-beam-type support structure (**Fig. 2**).

The second type, the semimonocoque structural configuration in which the stringer-stiffened skin shells are supported by frames, is the most efficient for the large near-circular fuselage bending loads (**Fig. 3**). *See* AIRFRAME.

Flight regime requirements. Once the aircraft's configuration has been established and the external contours of the aircraft defined, the designer needs information on the flight regime requirements such as the flight loads, pressurization level, and structural temperature range, the last associated with either the flight regime or the installation of equipment or engines.

Flight loads. The requirements of flight loads applied to fuselages can be established by a static balance of lift, drag, and balancing tail loads along with the loads imposed by payload or equipment distributions. The loading diagram in **Fig. 4** illustrates the loads applied to the aircraft and how they impose simple beam loads on the fuselage. The flight load requirements vary in magnitude depending on whether they are introduced by aircraft maneuvers, gust loads, or landing loads, and each must be considered in order to design all the structural elements of the fuselage. Concentrated loads are introduced at the attachment points of wings, empennage, engines, or landing gear. There are also flight control, equipment, and passenger or pilot accommodations equipment attachment loads.

Pressurization. Once the structural elements have been designed for either inertia or equipment-induced loads, the other major load consideration is the pressurization requirements. Pressurization loads are generated by the differential between the cabin pressure employed to sustain a comfortable atmosphere for the pilot or passengers and the ambient

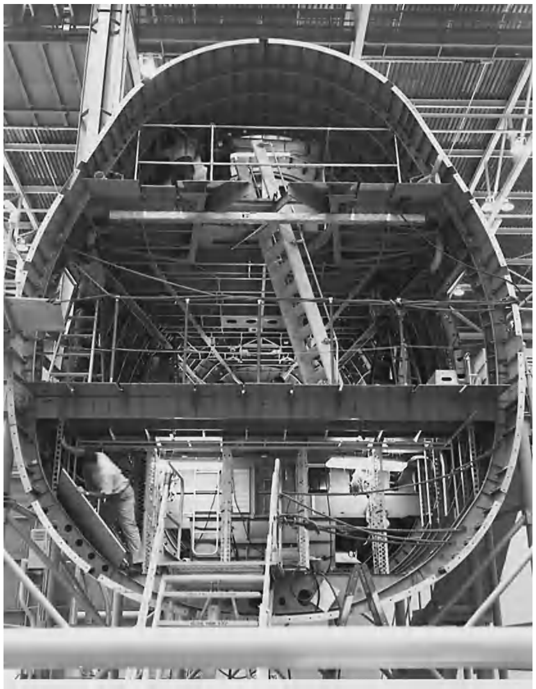


Fig. 3. Boeing 747 fuselage with stringer-stiffened skin supported by frames. (*Boeing Co.*)

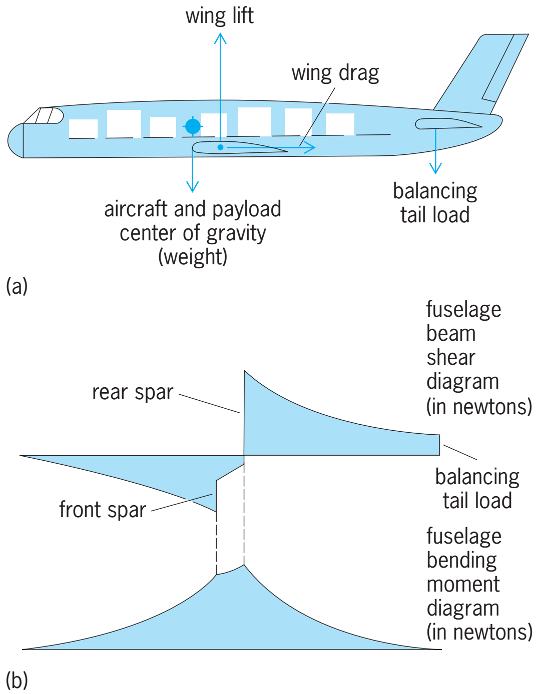


Fig. 4. Flight loads. (*a*) Aircraft static balance. (*b*) Fuselage simple beam loading.

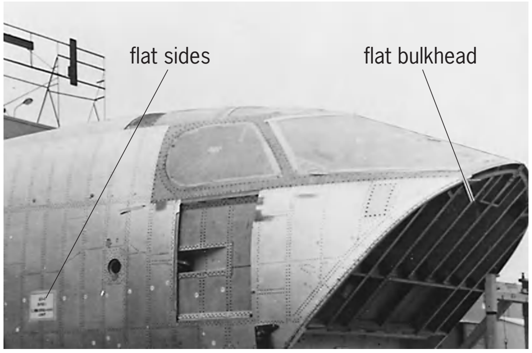


Fig. 5. Pressurized crew compartment of B-1 bomber.

altitude pressure associated with the flight regime. Since fighter and bomber fuselages are configurated for a minimum envelope and maximum aerodynamic efficiency, their shape may not be optimum for carrying pressure. A pressurized compartment containing flat sides may impose significant local internal structural loads since it is the most difficult structural configuration for sustaining pressure loads (**Fig. 5**). Fuselage configurations for large commercial transports are circular or near circular since this configuration is optimum for pressure containment. In the thin-shell fuselage structures this allows the skin to carry the pressure loads in hoop tension (**Fig. 6**) without introducing any significant loads in frames.

Temperature. The temperature requirements are associated with two potential sources of heat. One is the temperature change imposed on the structure by equipment such as engine exhaust, air-conditioning ducts, or armament discharge areas. The other structural temperature consideration is due to the ground environment and flight regime for either subsonic or supersonic performance requirements. The typical temperatures considered for the subsonic aircraft are those of ambient ground temperature plus solar heating and the −65°F (−54°C) temperature of high-altitude flight. For fighters and bombers and the British and French Concorde, the flight regime imposes aerodynamic heating temperatures up to 300°F (150°C). For this elevated-temperature supersonic flight regime, special design considerations must be taken into account when selecting the material to be used in the fuselage construction, as well as in the design of the structural details, to accommodate the temperature-induced deformation. *See* THERMAL STRESS.

Material selection. The above design requirements imposed by the flight and equipment loads, pressure containment, and the operational temperature range are all considerations in selection of the material to be used in a particular structural detail of the fuselage. The designer must consider the basic characteristics of those materials available for current aircraft construction. Aluminum, titanium, higher-strength steel, and some advanced composites such as graphite/epoxy have been used for construction of fighter and high-performance aircraft. A major portion of most cargo and commercial aircraft consists of aluminum, with highly loaded members, such as landing gear beams and engine mounts, of steel or titanium. Control surfaces of commercial aircraft are designed with graphite/epoxy materials. *See* COMPOSITE MATERIAL.

The basic structural characteristics of these materials are that aluminum provides a specific strength-to-weight density ratio on the order of 0.6×10^6 in. (1.5×10^4 m), while steels and titanium are of a similar strength-density ratio. The advanced composites

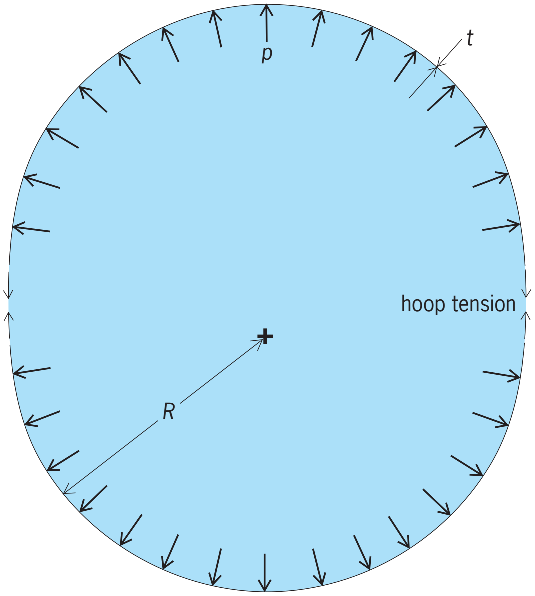


Fig. 6. Hoop tension from pressurization in thin skin shell. Here p = pressure, R = radius, t = skin thickness; hoop tension stress $f_t = pR/t$.

(graphite/epoxy or kevlar/epoxy) offer the advantages of higher strength-density ratios (1.2×10^6 in. or 3.0×10^4 m). *See* ELASTICITY; STRESS AND STRAIN.

Stability and strength. The configuration using the frame-stabilized major longeron and shear panel concept requires the designer to review each structural element for both strength and stability considerations. The frames provide the support, from a strength point of view, for all the local loads introduced by the attachment of wings, empennage, engines, landing gear, and armament, and the local loads induced by compartment pressurization. The longerons provide the main bending stiffness for the fuselage (Fig. 3). The major longerons act as columns between the frames; therefore, the stability of those columns must be a design consideration. *See* BEAM COLUMN.

In the fuselage structural configuration utilizing stringer-stiffened skin acting with frames, the frames stabilize or break up the skin stringer panel lengths. The stiffeners, with the skin, provide the structural stability and bending stiffness for the fuselage shell. This configuration is typical for commercial passenger and cargo aircraft (**Fig. 7**).

Private aircraft are often a combination of the above two types of structure. Major longerons and simple shear panels are used where large cutouts for passenger or cargo access are required. Stringer-stiffened skins are used for the continuous structure connecting the forebody to the empennage.

As discussed above, jet aircraft require structural elements which must support pressure loads. Those aircraft composed of major longerons and shear panels often have flat pressure-containment panels, which use diaphragm action or plate bending to support the pressure loads. In the circular fuselages, the hoop stresses induced in the shell provide pressure containment. Both of these pressure-containment configurations contribute to the fatigue and damage tolerance characteristics of the fuselage.

Fatigue. The fatigue considerations require a review of the repeated loads that are applied to the fuselage structure. These repeated loads are produced by the aircraft maneuvers, gusts, and typical landing and taxi loads. For pressurized areas, an additional requirement of the ground-air-ground pressure cycle must be introduced in the fatigue analysis. This ground-air-ground pressure cycle is often the dominant fatigue consideration.

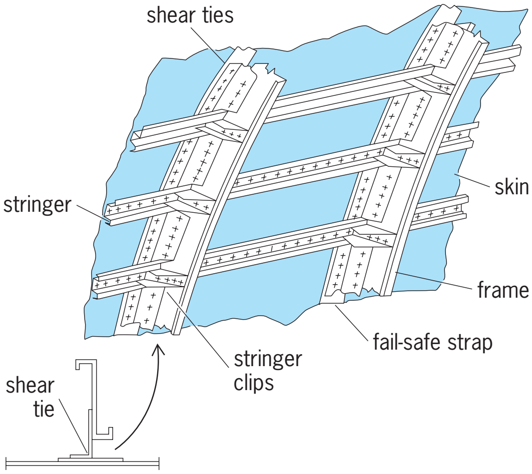


Fig. 7. Typical transport-aircraft aluminum fuselage panel.

Fatigue produced by this cyclic loading is most critical in the complex details (such as joints, splices, fittings, and cutout areas) where stress concentrations may occur at fastener holes, pad-up areas, and load discontinuities. The key considerations for fatigue are (1) the load type and its frequency, (2) the load-induced stress level at the structural detail, (3) the configuration of the detail or the magnitude of the stress concentration, (4) the fatigue characteristics of the material, and (5) the service life requirements. These life requirements vary dramatically, from the shorter usage life requirements of a fighter aircraft (about 6000 flight hours) to the long service life of a commercial passenger aircraft (about 60,000 flight hours). The magnitude of the loads encountered in the maneuver spectrum of a fighter aircraft are much higher than for commercial aircraft because of the fighter's combat maneuver requirements.

Once the elements of fatigue consideration have been assessed, a fatigue analysis is performed based on a damage accumulation rule. This damage accumulation rule is then used to assess the fatigue life of the aircraft.

Since many fatigue parameters are not finitely determinable, fatigue testing of the fuselage structural details and often of the total aircraft is part of the development of any aircraft fuselage. Since fatigue of metallic aircraft parts initiates crack growth in the material, this growth must be understood and its criticality assessed to establish the structure's damage tolerance *See* AIRCRAFT TESTING; METAL, MECHANICAL PROPERTIES OF.

Damage tolerance. Developments in fracture mechanics have provided design and analysis tools adequate to assure the damage tolerance of modern aircraft. Damage tolerance is the ability of an aircraft to tolerate fatigue-initiated flaws as well as external damage. This external damage may be due to such things as engine disintegration or foreign-object damage, and for fighter aircraft may be due to enemy weapons fire. This tolerance to damage provides the safety in modern commercial aircraft which makes them one of the safest means of transportation.

The approach to damage tolerance involves an understanding of the types of flaws that will occur, the rate at which these flaws will grow, the length at which the flaw will become critical for a specific material and structural configuration, and the length at which the flaw will become detectable. A fracture mechanics parameter describing the roughness of a material was developed, called the stress intensity factor. The application of this stress intensity factor allows the development of designs which can tolerate some predetermined-size flaw either by controlling the rate of crack growth or by providing a means of crack arrest. Whatever the means of containment, the designs must also take into consideration the ability to detect these flaws during inspections for maintenance of the aircraft. The inspection intervals and inspection methods must be established to ensure that a flaw is discovered prior to its reaching critical length.

The flaw arrest procedure generally has been used for the pressurized fuselages of commercial aircraft. Fail-safe straps or shear-tied frames are utilized to contain the crack in one or two bays between the fuselage frames (Fig. 7).

Other design considerations. While the structural requirements for strength, stability, fatigue durability, and damage tolerance are being met in the design of the fuselage, other configuration considerations must be taken in account by the designer, namely access to the interior of the aircraft or its contained equipment, the noise environment associated with the placement of the engines on, around, or in the aircraft fuselage, and maintenance and repair of the fuselage.

Interior access. In fighter fuselages, where the engines are buried in the main fuselage structure, the ability to access and remove these engines is a key consideration in the structure's functional design. In commercial aircraft, loading and unloading of the passengers and the passenger window accommodations are obviously very important.

Engine noise. Many commercial aircraft have engines mounted on the fuselage. The engine noise can induce sonic fatigue loads in the fuselage structural elements. The noise design requirements are met by reducing the noise level at the engine and by taking the noise loadings into consideration in the fatigue design of the fuselage. In addition, appropriate soundproofing is applied to the fuselage shell of commercial aircraft to limit the passenger annoyance.

Inspection, maintenance, and repair. Gross external damage must be detected during walkaround inspections or noted at the time of the occurrence during flight. Specially identified critical areas can be more thoroughly inspected by nondestructive inspection methods to estimate the extent of visible and nonvisible damage. These inspections will also cover those areas most likely to be damaged by mechanical process or atmospheric corrosion. The critical areas requiring special inspection instructions are identified at the time of the aircraft's design. Special inspections may include the removal of internal and external panels and a thorough inspection conducted with the aid of portable nondestructive inspection equipment, such as ultrasonics and x-ray. However, the most easily inspected designs provide visual access to both the external and internal sides of the fuselage shell. *See* NONDESTRUCTIVE EVALUATION.

The fuselage or its subassemblies are not removed from the aircraft for repair, as is often the case for ailerons, rudders, and elevators, or even in some cases empennages. Access to all parts of the fuselage structure for repair therefore is a necessity. Structural configuration compromises must be considered to facilitate repair and maintenance along with the other structural design requirements.

Weight and part reduction. For cargo and commercial aircraft the wing and fuselage are the two largest contributors to the aircraft structural weight, with each contributing 34–37% of the total structural weight. The fuselage weight of large aircraft has grown faster than the wing weight, which suggests that major opportunities for weight saving are in new fuselage design concepts (**Fig. 8***a*). Also because of wing bending-moment relief due to weight in the wing, greater payload improvement is achieved for fuselage weight reduction than for wing weight reduction (Fig. 8*b*).

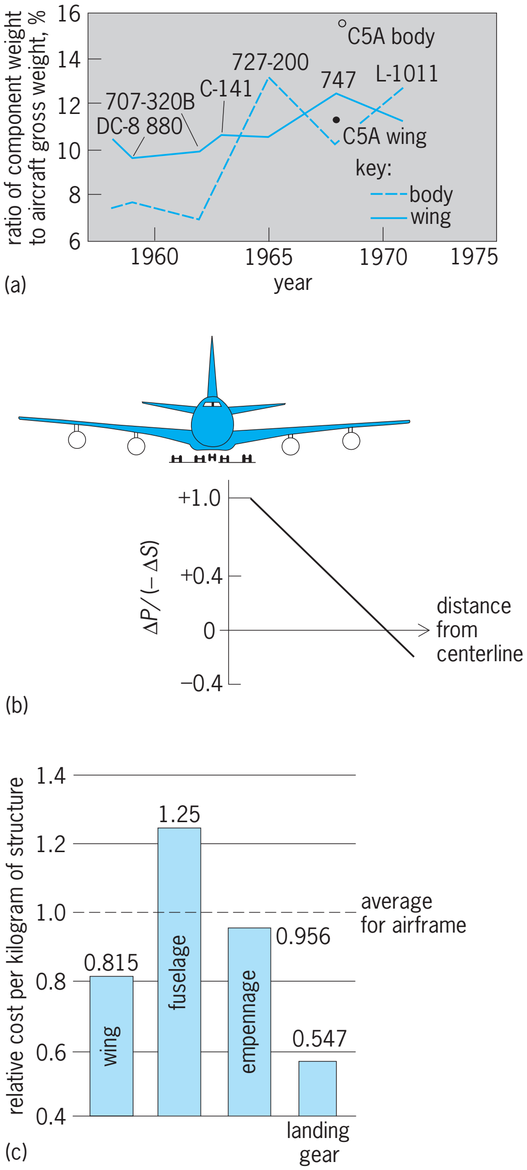


Fig. 8. Fuselage weight reduction. (*a*) Wing and fuselage weight history. (*b*) Plot of ratio of increase in payload ΔP to corresponding decrease in structural weight, $-\Delta S$, as a function of distance from aircraft centerline, showing centerline payload-weight effect. (*c*) Comparison of relative costs of aircraft components per kilogram of structure. (*After J. E. McCarty, Application of advanced structures and materials to transport aircraft fuselages, AIAA Pap. 74-340, AIAA/ASME/SAE 15th Structures, Structural Dynamics and Materials Conference, Las Vegas,1974*)

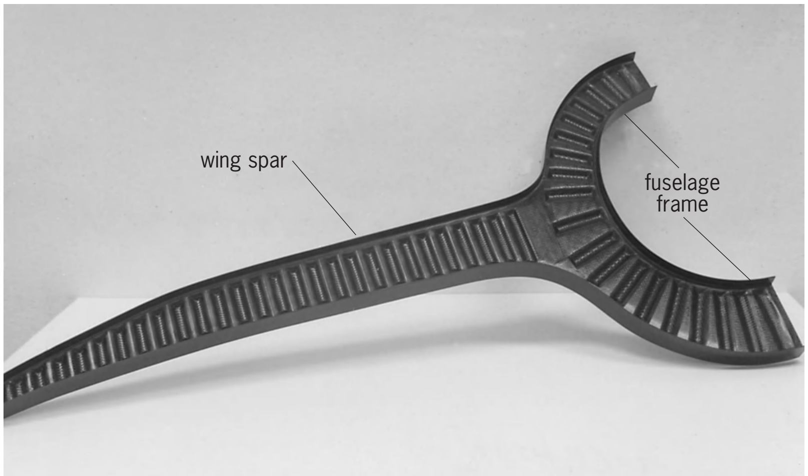


Fig. 9. Graphite/epoxy integral wing spar-fuselage frame for fighter aircraft. (*Boeing Research and Development*)

In addition to weight savings, Fig. 8*c* shows that the fabrication cost per pound of transport aircraft fuselage structure is 25% higher than the airframe average. Since the fuselage is approximately 37% of the airframe structural weight, high leverage in reducing airframe cost is available. New materials, configuration changes, and construction methods can reduce weight and cost of aircraft fuselages.

The material that dominates aircraft construction is aluminum; however, advanced composites are being introduced into empennage and wing structures of fighter aircraft, and advanced composites are to be applied to major aircraft fuselage structures. For circular fuselages, filament winding of composites may provide aircraft designers with the means to reduce weight and fabrication cost of large fuselages.

In addition to the application of advanced composite structures to aircraft fuselages, certain configuration adjustments are being made to high-performance fighter aircraft. Integral wing spar/frame designs allow more freedom of engine placement and better external aerodynamic contours, and provide a minimum weight configuration (**Fig. 9**). *See* WING STRUCTURE.

Fig. 10. Bonded honeycomb sandwich fuselage panel. (*After J. E. McCarty, Application of advanced structures and materials to transport aircraft fuselages, AIAA Pap. 74-340, AIAA/ASME/SAE 15th Structures, Structural Dynamics and Materials Conference, Las Vegas,1974*)

The application of adhesive-bonded aluminum fuselage structure in the configuration of bonded honeycomb sandwich construction can be utilized to improve the fatigue- and damage-tolerant characteristics of pressurized fuselages (**Fig. 10**). The adhesive-bonding assembly methods also reduce the construction cost of the fuselage, by reducing part count and mechanical fastening requirements. *See* AIRPLANE.

John E. McCarty

Bibliography. J. Cutler, *Understanding Aircraft Structures*, 3d ed., 1999; B. K. Donaldson, *Analysis of Aircraft Structures: An Introduction*, 1992; B. C. Hoskin and A. A. Baker (eds.), *Composite Materials for Aircraft Structures*, 1986; T. H. Megson, *Aircraft Structures for Engineering Students*, 3d ed., 1999.

Fusulinacea

An extinct group of marine organisms making up the superfamily Fusulinacea of the phylum Protozoa, order Foraminiferida. Fusulinaceans first appeared late in the Mississippian Period, and in the succeeding 100,000,000 years evolved into more than 125 genera and 6000 species before becoming extinct near the close of the Permian Period. Among protozoa fusulinaceans became giants, many reaching 0.4 in. (1 cm) in length and some even attaining 4 in. (10 cm) in length. The group is characterized by rapid evolution of distinctive morphological features, which have enabled paleontologists to establish a detailed fusulinacean phylogeny and biostratigraphic zonation. These fossils are widespread in Upper Mississippian, Pennsylvanian, and Permian marine strata in the major Eurasian and American geosynclinal belts and adjacent marine shelves, and have been reported from all continents except Australia and Antarctica, as indicated in **Fig. 1**.

Morphology. The structure of the families of the Fusulinacea is shown in **Fig. 2**. The calcareous shell of fusulinaceans is constructed on a simple plan. The initial chamber, the proloculus, is a small calcareous sphere with an opening on one side. As the individual grew and the volume of protoplasm increased, chambers were added successively in a planispiral coil around the proloculus. In most genera these chambers overlap the ends of the previous chambers and become elongate, parallel to the axis of coiling. The spiral wall (spirotheca), as in Fusulinidae, is composed of several layers. The initial wall is formed of a thin, dark organic-rich layer (tectum), beneath which is a thicker translucent layer (diaphanotheca) formed of calcium carbonate.

In several families of the Fusulinacea a secondary veneer of calcareous material (tectorial deposits) coats the interior of chambers and thus covers the floor, sides, or roof, and sometimes all three, of the chambers.

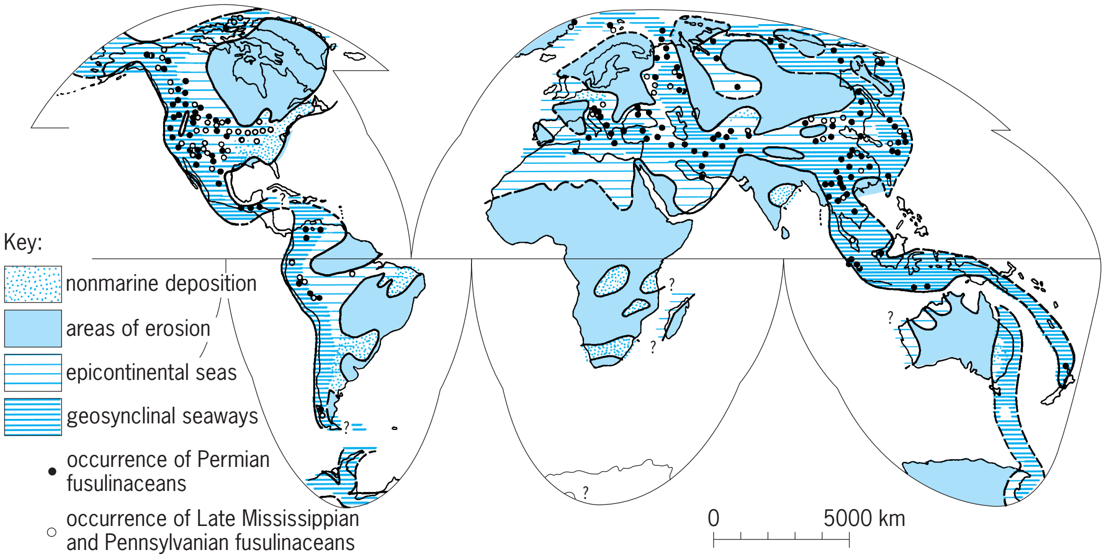


Fig. 1. Occurrences of Fusulinacea and distribution of sediments in the Pennsylvanian and Permian periods. 1 km = 0.6 mi.

Very small pores (mural pores), and 2 micrometers in diameter, penetrate the tectum and extend through the spiral wall. In highly evolved families, such as the Verbeekinidae and Schwagerinidae, a reduction in the number of layers in the spiral wall is accompanied by a joining together in the translucent layer of several mural pores to form larger tubes (alveoli), which gives the layer a honeycombed construction and modifies it to one distinctive layer, the keriotheca.

The frontal wall of each chamber forms a septum which divides the interior of the shell. In primitive genera septa are nearly planar, and only near the axial poles are they gently folded. In many of the more advanced genera the septa commonly become strongly folded across the entire length of the chamber. Larger pores (septal pores), about 10 μm in diameter, are distributed across the face of the septa.

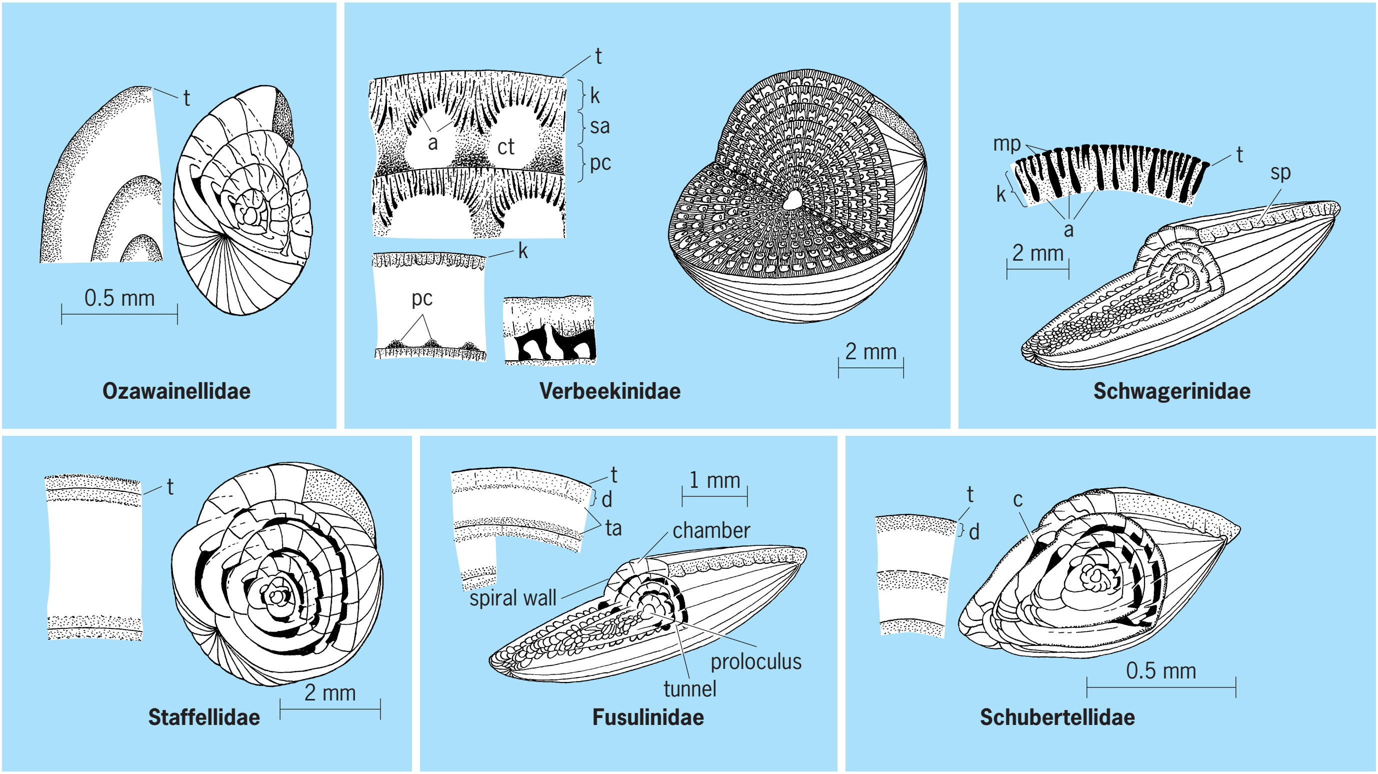


Fig. 2. The six families of Fusulinacea. Representatives shown in cutaway diagrams, with wall structures of each greatly enlarged; a = alveoli, c = chomata, ct = chamberlets, d = diaphanotheca, k = keriotheca, mp = mural pores, pc = parachomata, sa = septula, sp = septal pores, t = tectum, and ta = tectorial deposits.

Within the shell of fusulinaceans, a spiral tunnel is formed by resorption of part of the septa near the midplane. In many genera secondary deposits form ridges (chomata) along the floor of the chambers adjacent to the tunnel, as in Schubertellidae. This tunnel probably served as a passageway, along which the nucleus of the cell moved from the center of the shell toward the outer chambers. Most genera have single tunnels. A few have multiple or auxiliary tunnels and thus fusulinaceans may have developed multiple nuclei. In some advanced genera additional passageways between chambers (cuniculi) were formed by resorption of septa where opposing folds of adjacent septa meet. In the Verbeekinidae, secondary transverse and longitudinal partitions (septula) project downward from the spiral wall, and transverse

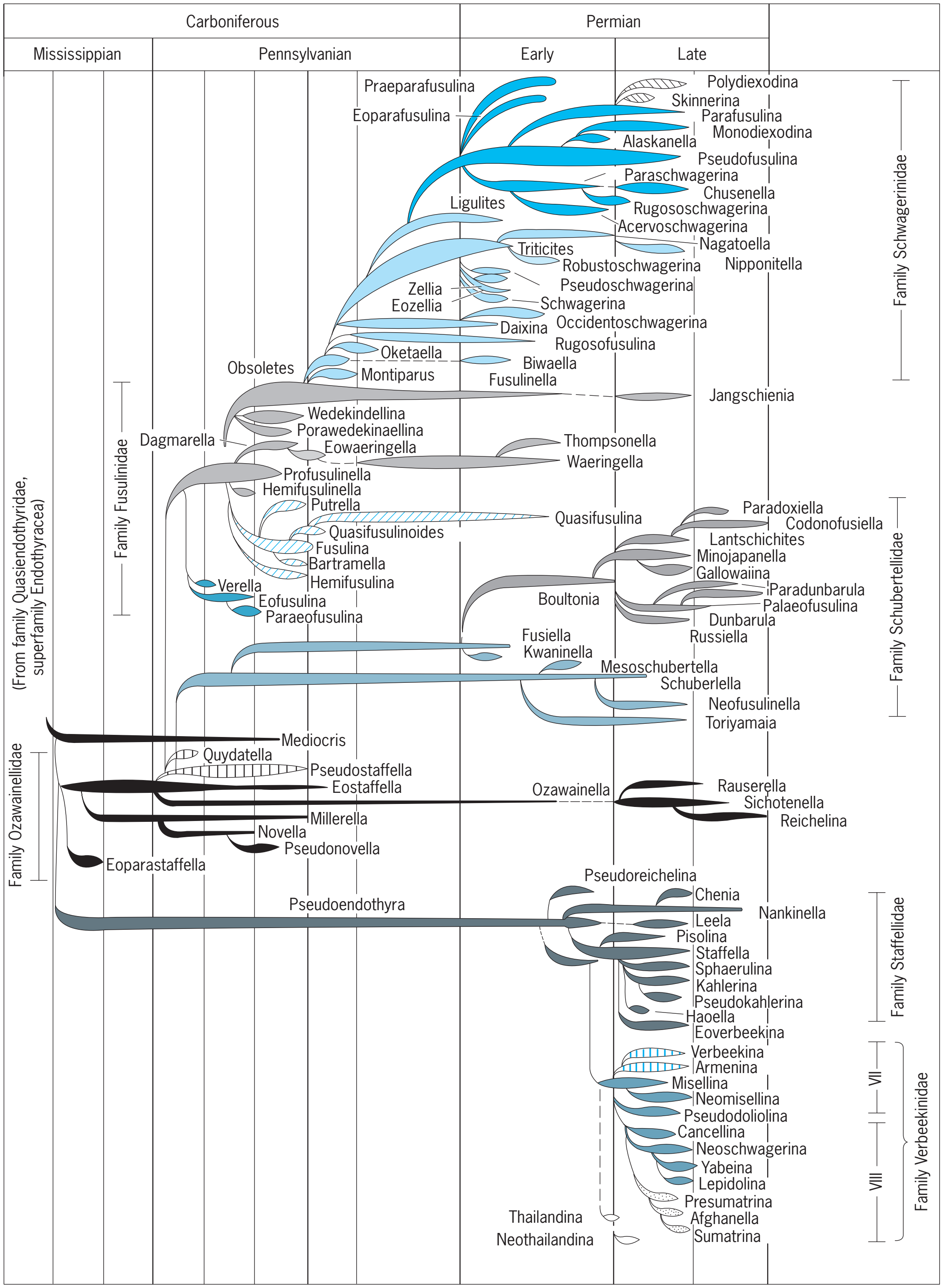


Fig. 3. **Evolutionary relationships of major genera and families of fusulinaceans.**

deposits (parachomata) form ridges on the floor of chambers to subdivide the chambers into chamberlets.

Evolution. Six evolutionary lineages of fusulinaceans (**Fig. 3**) corresponding to the six families, are differentiated on the basis of general shell structures and structural changes of the shell. The families Ozawainellidae, Schubertellidae, and Staffellidae persisted as small, simple, and apparently unspecialized lineages for most of the Pennsylvanian and Early Permian. During the Middle and Late Permian, new genera evolved in these three families and many of these genera were adapted to specialized ecological conditions, probably shallow, warm marine water having slightly above-normal salinity, conditions which appear to have persisted locally until near the end of the Permian Period.

The family Fusulinidae had great evolutionary diversification during the Early and Middle Pennsylvanian, and then within a remarkably short time came close to extinction before the beginning of the Late Pennsylvanian. Only a few genera of Fusulinidae survived, and this much-reduced lineage is traceable into the Late Permian.

The family Schwagerinidae arose from the Fusulinidae at about the end of the Middle Pennsylvanian, evolved slowly during the Late Pennsylvanian, and produced a burst of nearly 20 new genera during the Early Permian. The Schwagerinidae gradually, and then more rapidly, dwindled during the later part of the Early Permian, and finally became extinct near the middle of the Late Permian.

The sixth family, the Verbeekinidae, evolved from the Staffellidae in the Early Permian, underwent a rapid and distinctive evolutionary development, and with almost the same rapidity became extinct near the middle of the Late Permian. The Verbeekinidae and several genera in the other families had geographic distributions that were largely restricted to a Permian seaway that extended from northern Africa eastward through south-central Asia and southeastern Asia to New Zealand, and through China, Japan, and eastern Siberia into western North America as far south as California, but with rare exceptions not into other parts of Eurasia and North America and not into South America. *See* FORAMINIFERIDA. Charles A. Ross

Bibliography. K. Kanmera, K. Ishii, and R. Toriyama, *Geol. Paleontol. Southeast Asia*, 17:129–154, 1976; A. R. Loeblich and H. Tappan, *Treatise on Invertebrate Paleontology*, 2 vols., 1964; R. C. Moore, C. G. Lalicker, and A. G. Fischer, *Invertebrate Fossils*, 1952; C. A. Ross, *J. Paleontol.*, 41:1341–1354, 1967; S. E. Rozovskaya, *Trudy Paleontol. Inst. Akad. Nauk U.S.S.R.* (in Russian), vol. 149, 1975; R. R. Shrock and W. H. Twenhofel, *Principles of Invertebrate Paleontology*, 2d ed., 1953.

Fuzzy sets and systems

A fuzzy set is a generalized set to which objects can belong with various degrees (grades) of memberships over the interval [0,1]. In general, fuzziness describes objects or processes that are not amenable to precise definition or precise measurement. Thus, fuzzy processes can be defined as processes that are vaguely defined and have some cognitive uncertainty in their description. The data arising from fuzzy systems are, in general, soft, with no precise boundaries. Examples of such systems are large-scale engineering complex systems, social systems, economic systems, management systems, medical diagnostic processes, and human perception. *See* SET THEORY.

Function. The mathematics of fuzzy set theory deals with the uncertainty and fuzziness arising from interrelated humanistic cognitive types of phenomena such as subjectivity, thinking, reasoning, cognition, and perception. This type of uncertainty is characterized by structures that lack sharp (well-defined) boundaries. This approach provides a way to translate a linguistic model of the human thinking process into a mathematical framework for developing computer algorithms for computerized decision-making processes. The theory has grown very rapidly. Many fuzzy algorithms have been developed for application to process control, medical diagnosis, management sciences, engineering design, and many other decision-making processes where soft data are generated. Thus, fuzzy mathematics provides a modeling link between the human reasoning process, which is vague, and computers, which accept only precise data.

For example, in the design of many engineering systems, process information is not available both because it is difficult to understand precisely the complexity of the phenomena and because human reasoning is inexact and is based upon subjective perception. However, by virtue of knowledge and experience, which are inexact, it is possible to build increasingly good systems. In fact, fuzziness in thinking and reasoning processes is an asset since it makes it possible to convey a large amount of information with a very few words. However, in order to emulate this experience and these reasoning processes on a computer, for example, for intelligent robotics applications and medical diagnosis, a mathematical precision must be given to the vagueness of the information so that a computer can accept it. This is done by using the theory of fuzzy sets. Probability theory deals with the uncertainty or randomness that arises in mechanistic systems, whereas fuzzy set theory has been created to deal with the uncertainty that arises in human cognitive processes. *See* PROBABILITY.

The premise of fuzzy set theory is that the key elements in human reasoning processes are not numbers but labels of fuzzy sets. The degree of membership is specified by a number between 1 (full membership) and 0 (full nonmembership). An ordinary set is a special case of a fuzzy set, where the degree of membership is either 0 or 1. By virtue of fuzzy sets, human concepts like small, big, rich, old, very old, and beautiful can be translated into a form usable by computers.

Elements of fuzzy set theory. A fuzzy set A in the universe X is a set of ordered pairs $A = \{x,\mu_A(x)\}$, where x is a member of X ($x \in X$) and $\mu_A(x)$ is the grade of membership of x in A. Thus, μ_A is a function of X over the interval [0,1], as stated symbolically in Eq. (1). A fuzzy set is said to be normal if and only

$$\mu_A : X \rightarrow [0, 1] \tag{1}$$

if the maximum value of $\mu_A(x)$ [max $\mu_A(x)$] equals 1 (**Fig. 1**).

The calculus of the fuzzy sets is based on the following operations:

1. Inclusion. A fuzzy set A is said to be included in a fuzzy set B if Eq. (2) is valid.

$$\mu_A(x) \leq \mu_B(x) \qquad \text{for all } x \in X \tag{2}$$

2. Intersection. The intersection of the fuzzy sets A and B is defined by Eq. (3). The intersection corresponds to the logical connective AND. Thus, Eq. (4) holds.

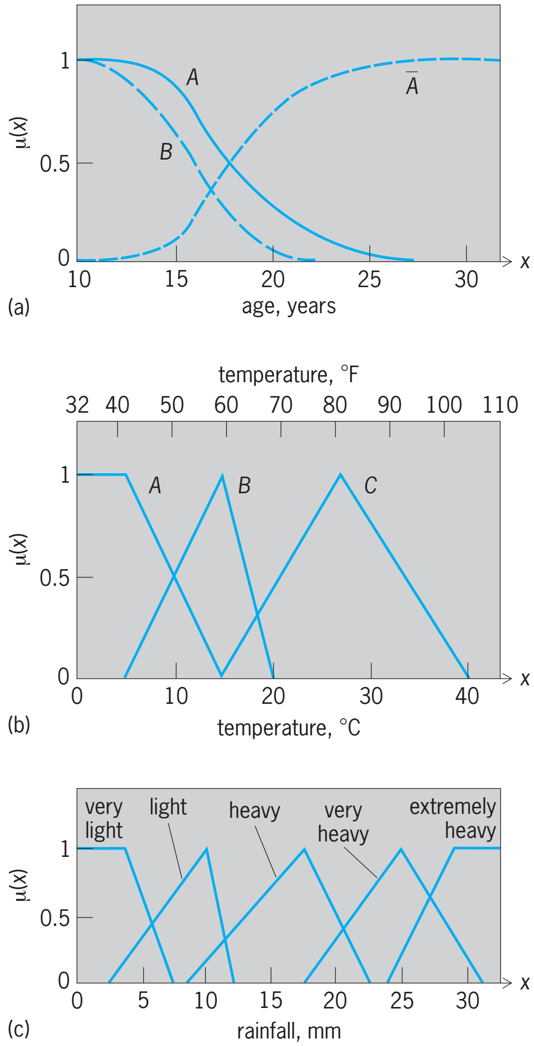


Fig. 1. Examples of fuzzy sets. (*a*) Fuzzy sets of age groups of high school children: young (*A*); not young ($\bar{A}$); very young (*B*). (*b*) Typical triangular representations of fuzzy labels of temperature: less than 6°C [43°F] (*A*); about 15°C [59°F] (*B*); and about 27°C [81°F] (*C*). These fuzzy labels may also be interpreted as: *A* = very cold; *B* = Cold; *C* = warm. (*c*) Graphs of membership used for classifying the rainfall in a region. 1 mm = 0.04 in.

$$\begin{aligned}\mu_{A\cap B}(x) &= \min[\mu_A(x), \mu_B(x)] \\ &= \mu_A(x) \wedge \mu_B(x)\end{aligned} \tag{3}$$

$$A \text{ AND } B = A \cap B \tag{4}$$

3. Union. The union of fuzzy sets A and B is defined by Eq. (5). The union corresponds to the connective OR. Thus, Eq. (6) follows.

$$\begin{aligned}\mu_{A\cup B}(x) &= \max[\mu_A(x), \mu_B(x)] \\ &= \mu_A(x) \vee \mu_B(x)\end{aligned} \tag{5}$$

$$A \text{ OR } B = A \cup B \tag{6}$$

4. Complement. The complement $\bar{A}$ of a fuzzy set A is defined by Eq. (7). The operation of comple-

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \tag{7}$$

mentation corresponds to the negation NOT. Thus, Eq. (8) holds. An example of this operation is shown in Fig. 1*a*.

$$\bar{A} = \text{NOT } A \tag{8}$$

5. Product. The product of two fuzzy sets is defined by Eq. (9).

$$\mu_{A\cdot B}(x) = \mu_A(x) \cdot \mu_B(x) \tag{9}$$

6. Algebraic sum. The algebraic sum is defined by Eq. (10).

$$\mu_{A+B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x) \tag{10}$$

In contrast to ordinary set theory, the union of a set and its complement is not necessarily the universe X; and the intersection of a set and its complement is not necessarily the null set ϕ. These two important properties are expressed in Eqs. (11) and (12).

$$A \cup \bar{A} \neq X \text{ (not necessarily)} \tag{11}$$

$$A \cap \bar{A} \neq \phi \text{ (not necessarily)} \tag{12}$$

An example that illustrates these operations is provided by two fuzzy sets A and B, defined over the universe X, where X is given by Eq. (13), and the sets

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6\} \tag{13}$$

A and B are defined by Eqs. (14) and (15), where numerical values over the interval [0,1] indicate the

$$\begin{aligned}A = \{&(0.6/x_1), (0.4/x_2), (0.3/x_3), (0.9/x_4), \\ &(1.0/x_5), (0.4/x_6)\}\end{aligned} \tag{14}$$

$$\begin{aligned}B = \{&(0.3/x_1), (0.8/x_2), (0.4/x_3), (0.7/x_4), \\ &(0.6/x_5), (1.0/x_6)\}\end{aligned} \tag{15}$$

grade of the membership μ. The OR, AND, and COMPLEMENT operations of the fuzzy sets A and B are given by Eqs. (16), (17), and (18) respectively.

$$A \cup B = \{(0.6/x_1), (0.8/x_2), (0.4/x_3), (0.9/x_4), (1/x_5), (1/x_6)\} \quad (16)$$

$$A \cap B = \{(0.3/x_1), (0.4/x_2), (0.3/x_3), (0.7/x_4), (0.6/x_5), (0.4/x_6)\} \quad (17)$$

$$\bar{A} = \{(0.4x_1), (0.6/x_2), (0.7/x_3), (0.1/x_4), (0/x_5), (0.6/x_6)\} \quad (18)$$

Other fuzzy set operations. The concept of partial membership makes possible other operations on fuzzy sets which are not found in the ordinary set theory. These operations are very important in modeling complex systems, for emulating the humanlike reasoning process in computers and in computer vision, and in modeling the human cognitive and perception process. Using abbreviated notation, some further important operations will be defined. *See* COGNITION; COMPUTER VISION; PERCEPTION.

The bounded product, bounded sum, and bounded difference of fuzzy sets *A* and *B* are defined by Eqs. (19), (20), and (21), respectively.

Bounded product:

$$\mu_{A \odot B} = O \vee (\mu_A + \mu_B - 1) \qquad (19)$$

Bounded sum:

$$\mu_{A \oplus B} = 1 \wedge (\mu_A + \mu_B) \qquad (20)$$

Bounded difference:

$$\mu_{A \ominus B} = O \vee (\mu_A - \mu_B) \qquad (21)$$

The concentration operation provides a relatively sharp boundary to a fuzzy set, and is defined by Eq. (22).

$$\mu_{\mathrm{CON}(A)} = \mu_A^{\alpha} \qquad \alpha > 1 \qquad (22)$$

The dilation operation provides a relatively flexible boundary to a fuzzy set, and is defined by Eq. (23).

$$\mu_{\mathrm{DIL}(A)} = \mu_A^{\beta} \qquad 0 < \beta < 1 \qquad (23)$$

The contrast intensification operation increases membership if it is greater than 0.5, and decreases if it is less than 0.5. It is defined by Eq. (24).

$$\mu_{\mathrm{INT}(A)} = \begin{cases} \mu_{\mathrm{CON}(A)} & \text{for } \mu(x) < 0.5 \\ \mu_{\mathrm{DIL}(A)} & \text{for } \mu(x) \geq 0.5 \end{cases} \qquad (24)$$

The blurring operation is the inverse of intensification and is defined by Eq. (25). Thus, the blurring operation, unlike the intensification operation, decreases membership if it is greater than 0.5, and increases it if less than 0.5.

$$\mu_{\mathrm{BLR}(A)} = \begin{cases} \mu_{\mathrm{DIL}(A)} & \text{for } \mu(x) < 0.5 \\ \mu_{\mathrm{CON}(A)} & \text{for } \mu(x) \geq 0.5 \end{cases} \qquad (25)$$

For the fuzzy set *A* of the above example, the membership vector is given by Eq. (26). Thus, for $\alpha = 2$

$$\mu_A = \{0.6, 0.4, 0.3, 0.9, 1.0, 0.4\} \qquad (26)$$

and $\beta = {}^1\!/_2$, the concentration, dilation, intensification, and blurring operations of the fuzzy set *A* are given by Eqs. (27)–(30).

$$\mu_{\mathrm{CON}(A)} = \mu_A^2 = \{0.36, 0.16, 0.09, 0.81, 1.0, 0.16\} \qquad (27)$$

$$\mu_{\mathrm{DIL}(A)} = \mu_A^{1/2} = \{0.77, 0.63, 0.55, 0.95, 1.0, 0.63\} \qquad (28)$$

$$\mu_{\mathrm{INT}(A)} = \{0.77, 0.16, 0.09, 0.95, 1.0, 0.16\} \qquad (29)$$

$$\mu_{\mathrm{BLR}(A)} = \{0.36, 0.63, 0.55, 0.81, 1.0, 0.63\} \qquad (30)$$

Systems modeling and control. Fuzzy set theory is a useful tool for modeling complex systems and for generating decision (control) rules. For example, a complex dynamic process with *x*(*t*) as an input and *y*(*t*) as an output (**Fig. 2**) can be modeled by using certain input and output observations. For example, a typical linguistic model would be given by Eq. (31),

$$\sum_i \text{IF } X_i \text{ THEN } Y_i \qquad (31)$$

which represents a collection of IF . . . THEN rules on labels of fuzzy input variables X_i and fuzzy output variables Y_i.

Control (decision) rules also can be generated for a feedback control process (**Fig. 3**) by using IF . . . THEN rules. The fuzzifier (F) converts the physical output *y*(*t*) into fuzzy variables Y_f. The fuzzy logic controller then generates the fuzzy control action U_f using IF . . . THEN rules given by Eq. (32). The defuzzifier (DF) converts the fuzzy control action U_f into a set of crisp values *u*(*t*) which in turn controls the dynamic system.

$$\text{IF } Y_f \text{ THEN } U_f \qquad (32)$$

Based upon such fuzzy control rules, several industrial controllers are in use for process control and similar operations. Also, very large scale integrated circuit (VLSI) hardware is being developed to implement these fuzzy logics. *See* CONTROL SYSTEMS; DECISION THEORY; INTEGRATED CIRCUITS.

Process control applications. A process operator, having an assigned control goal, observes the process state, controller, and process output, and intuitively assesses the variables and parameters of the process. From a subjective assessment of these quantities, the

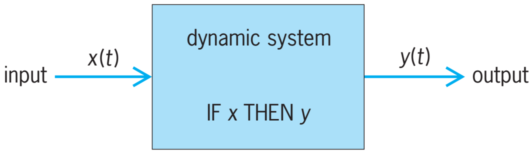


Fig. 2. Dynamic system with IF . . . THEN rules.

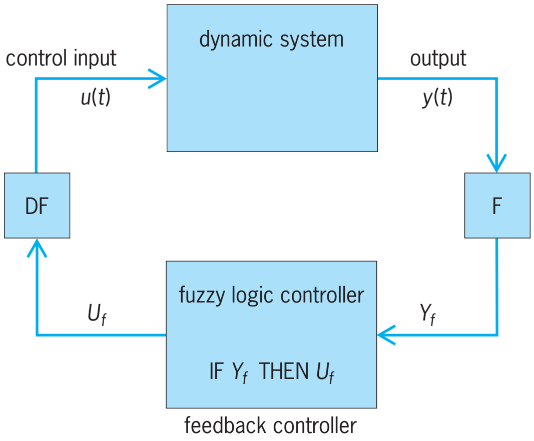


Fig. 3. Fuzzy logic controller using IF . . . THEN control rules. F = fuzzifier; DF = defuzzifier.

operator makes a decision and performs a manual alteration of the system control value so as to achieve an assigned control goal. It is possible to make a description of control strategy, a description of process behavior, and finally a subjective operational model of the operator's procedure. Such a control algorithm is very flexible and adequate to the actual process situation, and, as a rule, its performance is much better than that of algorithms obtained by using modern control theory. The linguistic algorithm comprises all the so-called metaphysical skills of the operator, such as intuition, experience, and intelligence, which cannot be dealt with by conventional mathematics. The operator's mind and knowledge of the controlled technological process can be formalized mathematically, to some extent, by using the theory of fuzzy sets and logic. *See* ALGORITHM; ARTIFICIAL INTELLIGENCE; EXPERT SYSTEMS.

For example, in observing an error E and error change ΔE, a process operator might carry out manual control of the process according to the following verbal description:

$$\text{If } E = \text{big and } \Delta E = \text{small}, \text{ then } U = \text{zero}$$

also

$$\text{If } E = \text{medium and } \Delta E = \text{zero}, \text{ then } U = \text{big}.$$

also

$$\vdots$$

In order to calculate control U, when given fuzzy relation R, error E, and error change ΔE, the compositional rule of inference given by Eq. (33) would be used, where $\circ$ is the max-min composition.

$$U = E \circ \Delta E \circ R \qquad (33)$$

There are four ways to obtain control rules: (1) they may be taken directly from the operator's experience; (2) they may be derived from a fuzzy model of the process, where the process characteristics are expressed in a form similar to control rules; (3) the operator's actions may be monitored and a control derived by using a fuzzy model of the operator's behavior; or (4) the rule may be learned by the controller.

Applications of fuzzy control theory include the fuzzy control of steam engines, heat exchangers, sinter plants, vehicular traffic, pressurized tanks, motors, ship autopilots, automobile speed, water cleaning processes, room temperature, pump operations, robots, prosthetic devices, warm-water plants, pressurized tank liquids, batch chemical reactors, nuclear fuel extraction processes, cement kilns, activated-sludge wastewater treatment, maintenance scheduling in transportation systems, stirred tanks, the process of changes of methane concentration, canvas production, fermentation processes of antibiotic synthesis, hydraulic servo systems, pulp plants, electric networks, converter steel-making processes, water-tank systems, robot arc welders, automatic train operation using predictive logic, automobile operation using self-learning, aircraft flight, and electrical trains, as well as fuzzy algorithms for path selection in autonomous vehicle navigation, and fuzzy adaptive control of continuous casting plants.

Fuzzy control of dc motor. An example of fuzzy control is provided by a direct-current (dc) motor (**Fig. 4**). The objective is to maintain constant motor speed despite the effect of a disturbing moment $M(t)$. The system may be described by the fuzzy variables S (error sum), I (motor current), and U (control action). The values of U corresponding to various values of S and I may be specified in a table

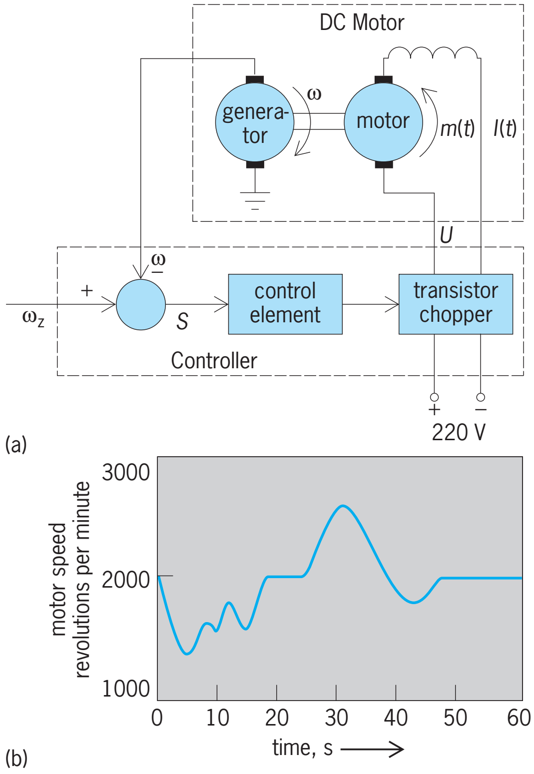


Fig. 4. Fuzzy control of a direct-current (dc) motor. (*a*) Elements of motor and controller. ω = motor angular speed, ω_z = reference angular speed, *M(t)* = disturbing moment, *S* = error sum, *I(t)* = motor current, *U* = control action. (*b*) Motor speed with fuzzy control.

Relation of fuzzy variables in control of a dc motor showing values of control action (*U*)

	Motor current (*I*)			
Error sum (*S*)	Zero	Small	Medium	Big
Super-big			Zero	
Big			Zero	Small
Medium		Null		Small
Small	Null	Null	Zero	Zero
Zero		Zero		Medium
Null		Zero		Big
Negative-null	Null		Zero	
Negative-zero	Null	Zero	Big	
Negative-small	Zero	Big		
Negative-medium	Null	Big	Big	
Negative-big	Zero	Big		
Negative-super-big	Small	Medium		

(see **table**) or in a list of relations, such as

$$\text{If } S = \text{null and } I = \text{small, then } U = \text{zero.}$$

These values satisfy the fuzzy relation determined by Eq. (34). The results of the experiment

$$\mu_R(s, i, u) = \vee\{\mu_S(s) \wedge \mu_I(i) \wedge \mu_U(u)\} \qquad (34)$$

(Fig. 4*b*) are very satisfactory. *See* ADAPTIVE CONTROL; INTELLIGENT MACHINE; PROCESS CONTROL; ROBOTICS. Madan M. Gupta

Bibliography. B. M. Ayyub and M. M. Gupta (eds.), *Uncertainty Analysis in Engineering and Sciences: Fuzzy Logic, Statistics, and Neural Networks Approach*, 1997; M. M. Gupta et al., *Approximate Reasoning in Expert Systems*, 1985; M. M. Gupta and N. K. Sinha, *Intelligent Control Systems*, 1996; M. M. Gupta and T. Yamakawa, *Fuzzy Computing: Theory, Hardware and Applications*, 1988; M. M. Gupta and T. Yamakawa (eds.), *Fuzzy Logic in Knowledge-Based Systems, Decision and Control*, 1988; A. Kandel and G. Langholz (eds.), *Fuzzy Control Systems*, 1993; A. Kaufmann and M. M. Gupta, *Fuzzy Mathematical Models in Engineering and Management Science*, 1988; A. Kaufmann and M. M. Gupta, *Introduction to Fuzzy Arithmetic: Theory and Applications*, 1985; R. Kruse, J. Gebhardt, and F. Klawonn, *Foundations of Fuzzy Systems*, 1994; W. Pedrycz, *Fuzzy Control and Fuzzy Systems*, 1993.

Fuzzy-structure acoustics

Large structures such as ships and airplanes can undergo a variety of complicated vibrations. Such structures typically consist of an outer body made of metal plating (for example, the hull of a ship) or perhaps a massive metallic frame (for example, the chassis of a truck), and a large variety of internal objects that are connected to either the plating or the frame. In designing such structures, it is highly desirable to have some method for predicting how they will vibrate under various conditions. The radiation of sound caused by these vibrations, either into the environment or into the empty portions of the structure, is also of interest because this sound is often either unwanted noise or a means of inferring information about the details of the structure or the excitation. Fuzzy-structure acoustics refers to a class of conceptual viewpoints in which precise computationally intensive models of the overall structure are replaced by nonprecise analytical models, for which the initial information is said to be fuzzy. *See* FUZZY SETS AND SYSTEMS.

External and internal structures. Fuzzy-structure theories divide the overall structure (**Fig. 1**) into a master structure and one or more attached structures, the latter being referred to as the fuzzy substructures, the internal structures, or the internals. (An example of a master structure is the hull and major framework of a ship.) The master structure is presumed to be sufficiently well known at the outset that its vibrations or dynamical response could be predicted if the forces that were exerted on it were known. Some of the forces are exerted on it by the substructures at the points at which they are attached. Such forces can be very complicated; nevertheless, there is some hope that a satisfactory approximate prediction of the vibrations of the master structure itself can be achieved with a highly simplified model.

The fuzzy substructures can be regarded as structures that are not known precisely. The most information-demanding type of fuzzy-structure model has significant drawbacks. Perhaps the most serious disadvantage is that the exploration of such models requires Monte Carlo simulations that are often infeasible or prohibitively time consuming. Consequently, there is some impetus to consider fuzzier models of the fuzzy substructures.

Forces and impedances. The fuzzier models that have been developed in the past few years rest on derivable consequences of the science of mechanics. The substructures are passive, so the displacements of attachment points on a substructure can be related to the forces exerted on them. The same relations, because of Newton's third law, have to hold

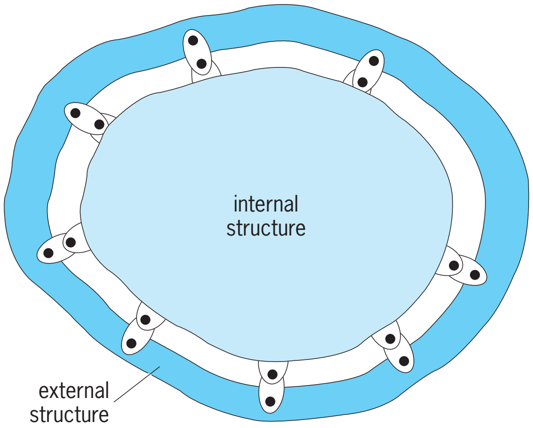


Fig. 1. Mechanical structure conceptually divided into an internal structure (fuzzy substructures) and an external structure (master structure).

between the forces exerted on the master structure and the corresponding displacements, except for a simple change in sign. The influences of the substructures on the master structure are analogous to those of springs connecting the master structure to rigid walls, but here the springs are more complicated than those described by Hooke's law (in which force is directly proportional to extension) in that displacements at distant attachment points may affect a given force, the forces are not necessarily in phase with the displacements, and the proportionality and the phase shift vary with the frequency with which the master structure is oscillating.

Nevertheless, there is still a surprisingly simple form for the general relations between the set of forces and the set of attachment-point displacements. Typically, workers in acoustics and vibrations express such relations as a rectangular array, any one element providing a ratio of a force experienced at one attachment point to the velocity of another attachment point, given that the displacements of all the other attachment points are held to zero, that the oscillations are sinusoidal (fixed frequency), and that complex numbers are used for the forces and velocities (so that differences in phases can be accounted for). Ratios of complex force amplitudes to complex velocity amplitudes are termed impedances, and an array of ratios of different forces to different velocities is termed an impedance matrix. *See* ACOUSTIC IMPEDANCE; MATRIX THEORY; MECHANICAL IMPEDANCE.

Modes and natural frequencies. The branch of mechanics known as linear vibration theory leads to the prediction that each of the impedance matrix elements can be written as a sum of terms, each term being similar to what would be expected from shaking a simple oscillator (**Fig. 2**), consisting of a mass, a spring, and a dashpot at an attachment point. Such a term, representing the ratio of a force to a velocity, has a frequency-dependent denominator that would go to zero in the absence of damping at an angular frequency (referred to as the resonance frequency or the natural frequency) equal to the square root of the spring constant divided by the mass. For very weak damping, the shaking force and the displacement are in the same direction when the shaking frequency is less than the natural frequency, but oppositely directed when the shaking frequency is higher than the natural frequency. *See* FORCED OSCILLATION; RESONANCE (ACOUSTICS AND MECHANICS).

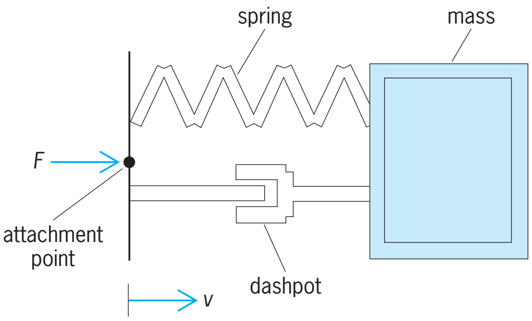


Fig. 2. Simple oscillator driven at its attachment point. *F* is the driving force, and *v* is the velocity at the attachment point. The ratio of the complex amplitude of *F* to the complex amplitude of *v* gives the driving-point impedance.

Typical substructures, when regarded in terms of the impedance matrix that they present to the external environment, behave as a set of independent simple oscillators, each having its own natural frequency. Corresponding terms in the series for the individual impedance matrix elements have the same denominator; consequently they go to zero, in the limit of zero damping, at the same natural frequency. These natural frequencies represent the frequencies at which the substructure would freely vibrate indefinitely in the absence of internal damping, when all of the attachment points are held motionless but the rest of the substructure is free. The overall manner in which the substructure is oscillating when undergoing vibrations under such circumstances is called a natural mode of vibration.

Modal masses. The numerators of the corresponding modal terms in the separate elements of an impedance matrix lead to the identification of a modal mass matrix for each natural mode. Each element of the modal mass matrix has the units of mass. For the special case where there is only one attachment point and the points in the substructure all move back and forth in the same direction, the matrix for each mode is just a one-by-one matrix, so there is a single mass associated with each mode. If there are a number of different attachment points but the points in the substructure still are constrained to move in the same direction, then the elements of the modal mass matrix for a given mode sum to a quantity identified as the modal mass for that mode. Finally, if substructures are considered where the points can move in all three directions and the definitions of the forces and velocity are generalized so that each cartesian component of each force is regarded as a separate force, then the sum of all the modal mass matrix elements for any given mode has a value that is three times as large as what would be regarded as the modal mass associated with that mode. In each such case, the modal masses when summed over all natural modes yield the total mass of the substructure.

Modal mass per unit frequency. Recently developed theories of fuzzy structures lead, after various plausible idealizations, to a formulation that requires only a single function, this being the modal mass per unit frequency bandwidth. Such a function could be regarded as a quasi-limit. A band of frequencies is chosen that is of narrow bandwidth but nevertheless contains a large number of natural frequencies. A sum is taken over all the modal masses for all of the modes that have frequencies in that band, this sum is then divided by the bandwidth, and the resulting value is associated with the center frequency of the band. This concept is similar to that of modal density, which is the number of resonance frequencies per unit frequency bandwidth—only here there is a weighting by modal mass.

This function can be estimated without the necessity of calculating each resonance frequency and

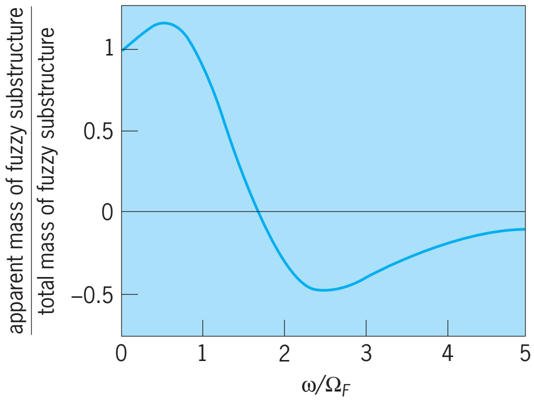


Fig. 3. Sample calculation of the frequency dependence of the apparent mass per unit area added to a master structure (rectangular plate) by a distribution of fuzzy structures when the master structure is oscillating at angular frequency ω. Ω_F is a characteristic frequency for the set of fuzzy substructures. ***(After A. D. Pierce, V. W. Sparrow, and D. A. Russell, Fundamental structural-acoustic idealizations for structures with fuzzy internals, J. Vibrat. Acous., 117:339–348, 1995)***

each modal mass separately, with very little detailed knowledge of the design, configuration, or material properties of the substructure itself. Examples considered so far suggest that, for any general category of substructure, a simple functional form can be identified for the dependence on frequency with the necessity to specify only a small number of parameters. A judicious sampling of blueprints or parts lists, in conjunction with a few simple field measurements, might be sufficient to estimate these parameters.

Predictions. In the formulation alluded to above, the influence of fuzzy substructures attached to the master structure tends to resemble that of an added frequency-dependent mass attached to the master structure in parallel with a frequency-dependent dashpot connecting the master structure to a hypothetical rigid wall. The added mass is a frequency-weighted integral over the modal mass per unit natural frequency, the weighting being such that the natural modes whose natural frequencies are less than the driving frequency have a positive contribution, while those for which the natural frequencies are greater than the driving frequency have a negative contribution. The master structure can seem to be less massive than it actually is (**Fig. 3**) when the bulk of the substructure mass is associated with resonant frequencies less than the excitation frequency.

The apparent dashpot constant, equal to the ratio of the apparent reacting force to the velocity at the attachment point, has a frequency-dependent value that is directly proportional to the mass per unit natural frequency bandwidth evaluated at the driving frequency, with an added multiplicative factor of the square of the driving frequency. Because the passive nature of the attached substructures precludes their acting as an energy source, the dashpot constant is always positive.

The prediction that the actual value of the dashpot constant is independent of the values of any parameters associated with energy dissipation (or damping) within the substructures is only approximate, but it does hold in the limit of sufficiently weak damping. There may be some relevant mechanism by which energy is absorbed within the substructure, but the actual amount of energy absorbed is asymptotically independent of the mechanism's strength. Similar phenomena, in which the amount of energy absorbed seems independent of the strength of the absorbing mechanism, occur in other branches of physics. For example, the energy loss at a shock front is dependent only on the pressure jump at the front in the limit when the shock front is thin and the absorption mechanism is weak. For the fuzzy-structure case, the evidence so far suggests that most internal structures of interest are sufficiently resilient that the asymptotic limit independent of damping mechanisms in the internals is of practical interest. This limit can be realized when (1) the number of resonances is very large and the resonance bandwidth of any given mode, although narrow compared with the resonance frequency itself, is sufficiently wide that there is considerable overlap between the resonance peaks of different modes; or (2) when the excitation of the master structure, while having a narrow bandwidth compared to the central frequency, is sufficiently wide that it includes a large number of natural frequencies of the substructure; or (3) when the force excitation is a short-duration pulse of nearly constant frequency, but the duration of the pulse is short compared to the reverberation time of the substructure, the latter being roughly equal to the reciprocal of the bandwidth of a resonant frequency of the substructure.

One implication of the newly emerging fuzzy-structure theories is that, insofar as there is concern with the vibrations of only the master structure, it is possible to drastically curtail the estimation or measurement of any parameters within the substructures that are associated with internal damping. *See* MECHANICAL VIBRATION; VIBRATION. Allan D. Pierce

Bibliography. A. D. Pierce, Resonant-frequency-distribution of internal mass inferred from mechanical impedance matrices, with application to fuzzy structure theory, *J. Vibrat. Acous.*, 119:324–333, 1997; A. D. Pierce, V. W. Sparrow, and D. A. Russell, Fundamental structural-acoustic idealizations for structures with fuzzy internals, *J. Vibrat. Acous.*, 117:339–348, 1995; C. Soize, A model and numerical method in the medium frequency range for vibroacoustic predictions using the theory of structural fuzzy, *J. Acous. Soc. Amer.*, 94:849–865, 1993; M. Strasberg and D. Feit, Vibration damping of large structures induced by attached small resonant structures, *J. Acous. Soc. Amer.*, 99:335–344, 1996.

G

Gabbro — Gentianales

Gabbro

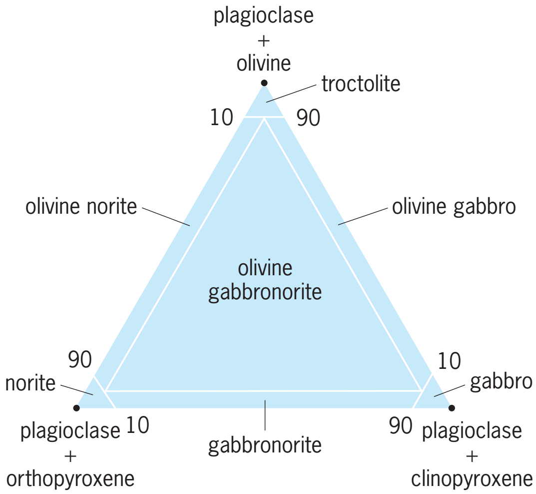
The plutonic equivalent of its more abundant extrusive equivalent, basalt. Because it crystallized from a magma intruded deep within the crust, gabbro has a grain size visible to the naked eye with approximately equal amounts of calcic plagioclase (with 50% or more anorthite, the calcium aluminium feldspar) and pyroxene. Olivine is common as an early crystallized mineral, but either nepheline or quartz could be a late-stage crystallization product found in the matrix. Hornblende or biotite is commonly formed as an alteration product of pyroxene during the late stages of the magmatic crystallization, when water becomes enriched in the residual magma. *See* BIOTITE; HORNBLENDE.

Gabbro is found in diverse tectonic environments, ranging through oceanic ridges, convergent plate boundaries, stable continents, and rifts. The forms of the intrusive gabbro bodies include dikes, sills, pipes, laccoliths, stocks, batholiths, and large layered intrusive complexes.

Nomenclature. The detailed terminology of gabbro is complex, involving two triangular diagrams and many varietal names; however, it can be simplified by utilizing one triangular diagram with three end members: calcic plagioclase and clinopyroxene, calcic plagioclase and orthopyroxene, and calcic plagioclase and olivine (see **illus.**). *See* FELDSPAR; OLIVINE; ORTHORHOMBIC PYROXENE; PYROXENE.

If the rock contains any hornblende or nepheline or quartz, these minerals become the modifier for the rock name. For example, if quartz is found with labradorite and clinopyroxene, the rock is called a quartz-bearing gabbro or a quartz gabbro. It is rare, however, to find quartz in gabbro containing visible magnesium-rich olivine because a chemical reaction takes place between the two to form the equilibrium mineral pyroxene. Thus, the term quartz gabbro is not used for a gabbro containing magnesium-rich olivine. In the rare cases where quartz is found with magnesium-rich olivine, the quartz usually will be rimmed with pyroxene, indicating that it was picked up accidentally by the gabbro magma and that it is clearly not in equilibrium. However, nepheline is stable with olivine. Hence, nepheline gabbro or nepheline troctolite, though rare, is not unusual. Nepheline and orthopyroxene, however, are mutually exclusive for chemical reasons again; hence, nepheline gabbronorite is not an appropriate term. *See* LABRADORITE; NEPHELINE; QUARTZ.

If the proportion of any of the three main phases increases over the other, new names are utilized. When the pyroxene exceeds plagioclase by a ratio of more than 7:3, the rock is called a melanogabbro or a melanonorite. When plagioclase is less than 10% of the rock, the rock is no longer a gabbro, and a classification scheme for ultramafic rocks must be adopted. However, when plagioclase exceeds 90% of the rock, the felsic (plagioclase feldspar–rich) rock is called an anorthosite. In the following discussion,



Triangular diagram showing end members and the nomenclature of gabbroic rocks.

gabbro will be used in a general sense to represent all seven names. *See* ANORTHOSITE.

Texture and structure. The grain size of gabbroic rocks ranges from a millimeter to centimeters. The finer-gained gabbro is commonly referred to as a diabase (or dolerite in the United Kingdom) that usually has small granular pyroxene interstitially enclosed by randomly oriented laths of calcic plagioclase. This diabasic texture results from the faster cooling of the magma due to its injection as small dikes and sills in shallow crust. In coarser-grained gabbro found in large plutons injected in deeper crust, the pyroxenes are larger and enclose partially (a texture called subophitic) or fully (a texture called ophitic) randomly oriented labradorite. Gabbros with mineral gains larger than 2 cm (0.8 in.) are rare, but such rocks are referred to as gabbro pegmatites. At contacts with country rock, the gabbro is commonly very fine grained or glassy because of the fast chilling of the magma. In some plutons, especially near their margins, gabbroic minerals are aligned perpendicular to the walls, yielding comb layering. *See* DOLERITE; MAGMA; PLUTON.

Chemical composition. Except for volatile compounds such as water, carbon dioxide, and sulfur molecules, which are largely lost during the eruption of basaltic magma, gabbros are chemically identical to basalts. Their silica (SiO_2) contents range 45–53% by weight, with 15–17% alumina (Al_2O_3). Gabbros have a high content of magnesium oxide (MgO), calcium oxide (CaO), and iron oxide (FeO) commonly ranging about 7–10 wt %, but the soda (Na_2O) and potash (K_2O) usually total less than about 3–4 wt % with the soda/potash ratio averaging less than 3. As with the basalts, some gabbros are alkali poor in relation to their silica content. As a consequence, saturated (with respect to silica) gabbros are specifically gabbronorites with two pyroxenes (both ortho and clinopyroxenes) and possibly quartz. These lack the undersaturated minerals such as nepheline. This variety is most common in the layered intrusions. The alkali-rich gabbros are undersaturated with respect to silica and, consequently, have only one pyroxene (clinopyroxene) and can have nepheline as an associated mineral.

Occurrence and origin. Unlike basalts, which are found in surface or near-surface environments, gabbros are found as shallow-to-deep intrusive bodies. Large blocks measuring more than 10 km (6 mi) thick, with a suite of rocks including serpentinite, pillow lavas, and chert, contain gabbro. This suite, called ophiolites, is commonly found on continents along convergent plate boundaries; it is thought to have been tectonically emplaced by thrusting onto the continental margins. It consists of a suite of rocks that are believed to represent the oceanic crust and upper mantle. This suite, exposed in places such as Newfoundland in Canada, Point Sal in California, Oman, and Cyprus (Troodos Massif), has a characteristic stratigraphy with the following sequence. From presumed top to bottom, there is an upper layer of oceanic sedimentary rocks (cherts) overlying basaltic pillow lavas that were fed from below by a diabasic sheeted dike complex. These rocks are above a layered gabbro sequence that overlies layered ultramafic rocks. Finally, at the bottom is an unlayered and metamorphosed ultramafic rock (serpentinite) representing the upper mantle. The sequence from the lower layered gabbro and ultramafic rocks to the sheeted dikes and the pillow lavas is believed to have resulted from fractional crystallization of a basaltic magma formed by partial melting of mantle rock. *See* CHERT; LAVA; OPHIOLITE; SERPENTINITE; STRATIGRAPHY.

Gabbros occur in minor abundance in the large granite batholiths found at convergent plate boundaries along the western Cordillera of the United States, including the Southern California, Sierra Nevada, Idaho/Boulder, and Coast Range (British Columbia) batholiths. The most extensive occurrence of gabbro (14% of total exposed area) is found in the Southern California batholith. Such occurrences are thought to be formed by fractional crystallization of the dominant granodioritic parent magma.

In Precambrian terranes, large batholiths (called massifs) of anorthosites (an igneous rock with greater than 90% plagioclase) commonly have gabbronorite at their margins. Such an occurrence is most common in the Grenville tectonic province (1000 to 900 million-year-old rocks) of eastern Canada and in Norway. The anorthosites and the gabbronorites are believed to have been formed by igneous processes and emplaced as a crystal-liquid mush into shallower regions of the crust.

Large layered intrusions such as the Stillwater Complex in Montana, the Skaergaard Intrusion in Greenland, the Muskox in Northwest Territories of Canada, and the one in the Bushveld Complex in South Africa have a layered gabbro/norite sequence similar to that found in the ophiolite suites. However, the homogeneous lower layer of serpentinite, the sheeted dike complex, the pillow lavas, and the chert are missing, and are replaced by a zone of silica-rich granophyre (quartz enclosed in sodic and potash feldspar, plus iron-rich olivine and pyroxene rock). As in ophiolites, an ultramafic layered sequence underlies the layered gabbro/norite sequence. The gabbronorite layers are continuous for hundreds of meters, yet are only a few centimeters to meters thick and alternate commonly with anorthosite layers. Within this layer the compositions of the pyroxenes and plagioclases become respectively iron-rich and sodium-rich from the bottom to the top. This layering is thought to originate in a manner similar to that in an ophiolite sequence by a fractionation process involving the settling of the denser grains from the liquid magma. *See* BASALT; IGNEOUS ROCKS.

Albert M. Kudo

Bibliography. A. Hall, *Igneous Petrology*, 2d ed., 1996; D. W. Hyndman, *Petrology of Igneous and Metamorphic Rocks*, 2d ed., 1985; A. R. McBirney, *Igneous Petrology*, 2d ed., 1993; A. R. Philpotts, *Principles of Igneous and Metamorphic Petrology*, 1990.

Gadiformes

An order of fishes containing the cods, codlings, codlets, grenadiers, and hakes. The Gadiformes plus the orders Ophidiiformes (pearlfishes, cusk-eels, and brotulas), Batrachoidiformes (toadfishes), and Lophiiformes (anglerfishes), collectively called the Anacanthini, are thought by some authors to be a monophyletic lineage but the concept is doubted by others. The group is characterized by the absence of a myodome (a cavity in the postorbital region of the skull in which lodge the muscles of the eye; actinopterygian fishes typically have a well-developed myodome); absence of parapophyses (singular parapophysis; a long, transverse process arising from the abdominal vertebral centrum, which serves to support epipleural ribs and, in Gadidae, the gas bladder; also called transverse process) on at least the first three vertebrae; and insertion of the first few pairs of ribs in cavities of the vertebral centra rather than on the parapophyses.

General characteristics. Members of the Gadiformes share the following characteristics: The pelvic fins are inserted below (thoracic) or anterior to (jugular) the pectoral fins, but rarely behind the pectoral fin (in some species in the family Macrouridae); there are no true spines in the fins; the dorsal and anal fins are long, with one, two, or three lobes; the scales are usually cycloid, rarely ctenoid; the premaxilla, which is protractile in some species, forms the entire margin of the upper jaw; the orbitosphenoid and basisphenoid bones are absent; the branchiostegal rays number six to eight; and the swim bladder is physoclistic. Some of the gadiforms have a complex glycoprotein antifreeze that protect them from internal ice formation at subzero water temperatures.

Classification and distribution. Concepts of the composition of Gadiformes have varied over the years. An early classification recognized only four families: Muraenolepididae, Moridae, Bregmacerotidae, and Gadidae. Then more families were recognized, including the Ophidiidae, Carapidae, and Zoarcidae, of which the first two are now placed in the order Ophidiiformes and the last in Perciformes. Currently 9 families, 75 genera, and 555 species are recognized, only one of which is limited to freshwater. About one-quarter of the world's commercial marine fishery is composed of gadiform fishes, primarily cod, haddock, pollock, and hake. In North American waters there are six families (Macrouridae, Moridae, Bregmacerotidae, Phycidae, Merluccidae, Gadidae) and 50 species.

Euclichthyidae (eucla cod). This is a monotypic family. The species *Euclichthys polynemus* is benthopelagic off New Zealand and Australia. The second dorsal and anal fins are continuous and not indented; pelvic fins are jugular, each of four long free rays, the longest reaching posteriorly to the anus; no chin barbel; maximum length 35 cm (14 in.).

Macrouridae (grenadiers). The grenadiers are classified in four subfamilies, 27 genera, and 350 species, which range in deep water from the Arctic to the Antarctic. Most of the species are benthopelagic, at depths between 200 and 2000 m (660 and 6600 ft), in tropical and subtropical waters. The second dorsal and anal fins are long and low and continuous with the tail, which tapers to a sharp point; the pelvic fins are thoracic but are behind the pectoral fins in some species; the caudal fin is absent as well as most of the caudal skeleton that supports it; the scales are small; a light organ may be present on the midline of the abdomen just in front of the anus; maximum length is 80 cm (32 in.).

In the subfamily Bathygadinae the second dorsal fin rays are longer than the anal rays; the mouth is wide and terminal; a chin is barbel present or absent; the snout is rounded; and the maximum length is 65 cm (26 in.). These fishes occur primarily from 200 to 2700 m (660 to 8900 m) in tropical to subtropical seas, but they are absent from the eastern Pacific.

The subfamily Macrouroidinae consists of two monotypic genera, one lacking a pelvic fin, the other with a small pelvic fin of five rays. The head is large, the mouth subterminal, the single dorsal fin is low; chin barbel absent; and the maximum length is 40 cm (16 in.). They are bentho- to bathypelagic fishes that occur worldwide in tropical to subtropical sea.

The subfamily Trachyrincinae consists of two genera, one lacking a chin barbel, the other with a small barbel. The second dorsal fin rays are somewhat longer than the anal fin rays, the mouth is wide and subterminal, the snout is long and pointed, the body scales are spinous, and the maximum length is 60 cm (24 in.). They occur practically worldwide in temperate waters.

The subfamily Macrourinae comprises 32 genera and 255 described species plus a number of undescribed species. The second dorsal fin rays are much shorter than those of the anal fin, the mouth is terminal to subterminal, a barbel is present or absent, many species have a ventral light organ, and the maximum length is 150 cm (59 in.) in one species but much less in the other species. They are practically worldwide but absent in high Arctic latitudes.

Moridae (morid cods). This family comprises 18 genera and about 105 species, which occur worldwide in deep waters. Morid cods have one or two, rarely three, dorsal fins and one or two anal fins; a chin barbel is present or absent; the swim bladder is connected with the auditory capsule (called otophysic connection); and some species have a light organ (see **illustration**).

Melanonidae (pelagic cods). Only two species in one genus, *Melanonus*, comprise this bathypelagic family known from the Atlantic, Indian, and Pacific oceans, as well as subantarctic waters. They have a single, long-based dorsal fin, reaching near but free from a small caudal fin; a barbel is absent; and there is no otophysic connection.

Bregmacerotidae (codlets). The one genus and 15 (probably more) species of this family are marine (rarely estuarine) fishes of tropical and subtropical seas. The first dorsal fin, consisting of a single elongate ray, is on the nape and separate from the second dorsal by

Beardie (*Lotella physcis*), a species of morid cod. (*Photo © John E. Randall*)

a long gap. The second dorsal and anal fins, long and with a wide notch in the middle, are mirror images; the pelvic fins are jugular, with five rays, the outer three of which are free filaments; and the maximum length is 12 cm (4.7 in.).

Muraenolepididae (eel cods). The four species of eel cods, in a single genus, *Muraenolepis*, occur primarily in Antarctic waters. The first dorsal fin has only two rays, the second dorsal and anal fins are continuous with the caudal fin; the pelvic fins are inserted under the gill openings, which extend upward just to the base of the pectoral fins; a chin barbel is present; and the maximum size is 40 cm (16 in.).

Phycidae (phycid hakes). Two subfamilies are recognized, Gaidropsarinae (rocklings) and Phycinae (phycine hakes). The rocklings have three closely set dorsal fins, the first a single ray, the second of low rays in a fleshy cover, and the third long and continuous to near the caudal fin; the anal fin is single and similar to the third dorsal; the snout has two to four prominent barbels plus additional rudimentary barbels in some species, and a single barbel is on the chin. The phycine hakes differ in having two dorsal fins, the first relatively short, the second long and similar to a smaller single anal fin; and snout barbels are absent but a chin barbel is present. The rocklings, comprising of three genera and 15 species, are found primarily in the northern Atlantic and to a lesser extent in the Southern Hemisphere off New Zealand and Tasmania. The phycine hakes, in two genera and 10 species, are limited to the Atlantic, including the Gulf of Mexico and the Mediterranean Sea.

Merlucciidae (merluccid hakes). This family is represented by three subfamilies. The subfamily Merlucciinae (merluccid hakes), comprising 13 species in a single genus, *Merluccius*, occurs in the western and eastern Atlantic (including the Mediterranean and Black seas), off the southern tips of India and New Zealand, and in the eastern Pacific from British Columbia to the tip of South America. There are two dorsal fins, the second deeply notched, giving the appearance of three dorsal fins; the anal fin is almost a mirror image of the second dorsal; the mouth is large and terminal and equipped with long teeth; the pelvic fins are thoracic. Commonly applied names of various species of *Merluccius* are hake and whiting; however, whiting is also used for some other gadiforms, and the genus *Urophycis* in the family Phycidae are also called hake.

The subfamily Macruroninae (southern hakes) has seven extant species in three genera. These are Southern Hemisphere fishes, primarily off the southern shores of Africa, South America, New Zealand, and Australia. The body is elongate and strongly compressed (flattened laterally); the pectoral fins are high on the body, above the axis of the body; usually the upper jaw is equipped with two rows of teeth, with those in the outer row caninelike.

The subfamily Steindachneriinae (luminous hake) is monolytic. The one species, *Steindachneri argentea*, is unique among the gadiforms in having the anus displaced forward to between the pelvic fins, leaving the urogenital pores in their normal place just in front of the anal fin. The caudal fin and skeleton are extremely reduced and the tail tapers to a fine point, and an elaborate light organ system is present on the head. The luminous hake occurs in the Gulf of Mexico, Caribbean Sea, and the Guianan shelf off northeastern South America.

Gadidae (cods). Three subfamilies are usually recognized, although many authors rank the subfamilies as families. A unifying character is a swim bladder that is not connected with the auditory capsule.

The subfamily Lotinae (cuskfishes) comprises three genera and five species. One species, the burbot (*Lota lota*), is limited to freshwater. It inhabits large deep rivers and lakes in the northern United States and throughout Canada and northern Eurasia. Some freshwater populations of another gadid, the Atlantic tomcod (*Microgadus tomcod*) of the Atlantic coast of North America, are known. The remaining three species are strictly marine in the Arctic and northern Atlantic, including the Mediterranean Sea. Lorinae are characterized by one or two dorsal fins and one anal fin; always a chin barbel, never a barbel on the snout; a rounded caudal fin; and eggs with an oil globule.

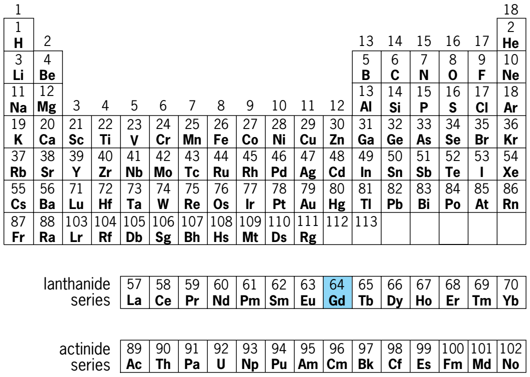
The subfamily Gadinae (cods and haddock), comprising 12 genera and about 25 species, occur in the Arctic, Atlantic, and Pacific oceans. These fishes have three dorsal fin and two anal fins, a chin barbel is usually present but no barbel on the snout; a truncate or slightly forked caudal fin; eggs without an oil globule; and maximum length 1.8 m (5.9 ft). Gadinae is considered the most advanced group in the order.

The Ranicipitidae (tadpole cod), a subfamily of Gadidae, is monotypic, with the single species *Raniceps rantus*, which occurs in shallow coastal waters of the northeastern Atlantic. The second dorsal and anal fins are continuous and not indented; pelvic fins are thoracic; a chin barbel is present; maximum length is 30 cm. Herbert Boschung

Bibliography. D. M. Cohen, Gadiformes: Overview, in H. G. Moser et al. (eds.), *Ontogeny and Systematics of Fishes*, Spec. Publ. no. 1, pp. 259–265, Amer. Soc. Ichthyol. Herp., 1984; D. M. Cohen (ed.), *Papers on the Systematics of Gadiform Fishes*, Nat. Hist. Mus. Los Angeles City, Sci. Ser. 32, 1989; D. M. Cohen et al., *FAO Species Catalogue*, vol. 10: *Gadiform Fishes of the World (Order Gadiformes): An Annotated and Illustrated Catalogue of Cods,* *Hakes, Grenadiers and Other Gadiform Fishes Known to Date*, FAO Fish. Synop. no. 10 (125), 1990; G. J. Howes, Anatomy, phylogeny and taxonomy of the gadoid fish genus *Macruronus* Günther, 1873, with a revised hypothesis of gadoid phylogeny, *Bull. Brit. Mus. Nat. Hist. (Zool.)*, 51(1):77–110, 1991; G. J. Howes, Biogeography of gadoid fishes, *J. Biogeog.*, 18:595–622, 1991; G. J. Howes, Phylogenetic relationships of macrouroid and gadoid fishes based cranial myology and arthrology, in D. M. Cohen (ed.), *Papers on the Systematics of Gadiform Fishes*, Nat. Hist. Mus. Los Angeles City, Sci. Ser. 32, pp. 113–128, 1989; G. J. Howes, The syncranial osteology of the southern eel-cod family Muraenolepididae with comments on its phylogenetic relationships and on the biogeography of sub-Antarctic gadoid fishes, *Zool. J. Linn. Soc.*, 100:73–100, 1990; D. F. Markle, Aspects of character homology and phylogeny of the Gadiformes, in D. M. Cohen (ed.), *Papers on the Systematics of Gadiform Fishes*, Nat. Hist. Mus. Los Angeles City, Sci. Ser., 32, pp. 59–88, 1989; J. S. Nelson, *Fishes of the World*, Wiley-Interscience, 1976; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, 2006; C. Patterson and D. E. Rosen, The Paracanthopterygii revisited: Order and disorder, in D. M. Cohen (ed.), *Papers on the Systematics of Gadiform Fishes*, Nat. Hist. Mus. Los Angeles City, Sci. Ser. 32, pp. 5–36, 1989.

Gadolinium

A metallic chemical element, Gd, atomic number 64 and atomic weight 157.25, belonging to the rare-earth group. The naturally occurring element is composed of eight isotopes. It is named in honor of the Swedish scientist J. Gadolin. The oxide, Gd_2O_3, in powdered form is white, and solutions of the salt are colorless. Gadolinium metal is paramagnetic and becomes strongly ferromagnetic below room temperatures. The Curie point, where this transition occurs, is about 16 K. *See* PERIODIC TABLE; RARE-EARTH ELEMENTS.



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner, Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Gages

Devices for determining the relative size or shape of objects. The function of gages is to determine whether parts are within or outside of the specified tolerances, which are expressed in a linear unit of measurement. Gages are the most widely used production tools for controlling linear dimensions during manufacture and for assuring interchangeability of finished parts. A gage may be an indicating type that measures the amount of deviation from a mean or basic dimension, or it may be a fixed type that simply accepts parts within tolerance and rejects parts outside tolerance. *See* TOLERANCE.

Gage tolerance. If parts are to be interchangeable, it is necessary that, when one device is used to control the accuracy of another, the controlling device itself ordinarily be accurate to the next significant figure. Thus, if a dimension of a part being manufactured is to be controlled accurately in thousandths of an inch, the gages to do this must be accurate in ten-thousandths of an inch. The gages must be set or measured by instruments accurate in hundred-thousandths of an inch. The instruments must be set by gage blocks accurate in millionths of an inch. The size of the gage blocks is verified by an interferometer reading in ten-millionths of an inch.

The tolerance allowed the gage maker and its effect on the size of the part can be illustrated as follows (**Fig. 1**). If the diameter of a 2.000-in. (5.080-cm) shaft is to be controlled during manufacture to 0.004-in. (0.01016-cm) tolerance, the "go" ring gage would have a wear allowance of 0.0002 in. (0.000508 cm) and a gage maker's tolerance of 0.0002 in. (0.000508 cm). Its size would be 1.9998 − 0.0002 in. (5.079492 − 0.000508 cm). The "not go" gage would have no wear allowance, but a gage maker's tolerance of 0.0002 in. (0.000508 cm). Its size would be 1.9960 + 0.0002 in. (5.06984 + 0.000508 cm). Thus, because of the necessity for a gage tolerance, if the part tolerance must be maintained absolutely, something less than the full part tolerance is available as a working tolerance (in this case 85–95%), depending upon the actual sizes of the "go" and "not go" gages.

Gage blocks. The standard of lineal measurement for most manufacturing processes is the gage block (**Fig. 2**). Gage blocks are used for setting gages, for setting machines, and for comparative measurements.

Gage blocks are made of chrome steel and may be chrome-plated or have carbide wear faces. Individual blocks have two surfaces which are flat and parallel. The parallel distance between the surfaces is the size marked on the block to a guaranteed accuracy of 0.000002, 0.000005, or 0.000008 in. (0.0508, 0.125, or 0.203 micrometer) at an ambient temperature of 68°F (20°C). Blocks are sold in sets so that combinations of individual blocks, when wrung firmly together, will produce a new end standard equal to the sum of the sizes of the individual blocks. With a set of 81 blocks, any length expressed in four decimal places from 0.2000 in. (0.5080 cm) to approximately

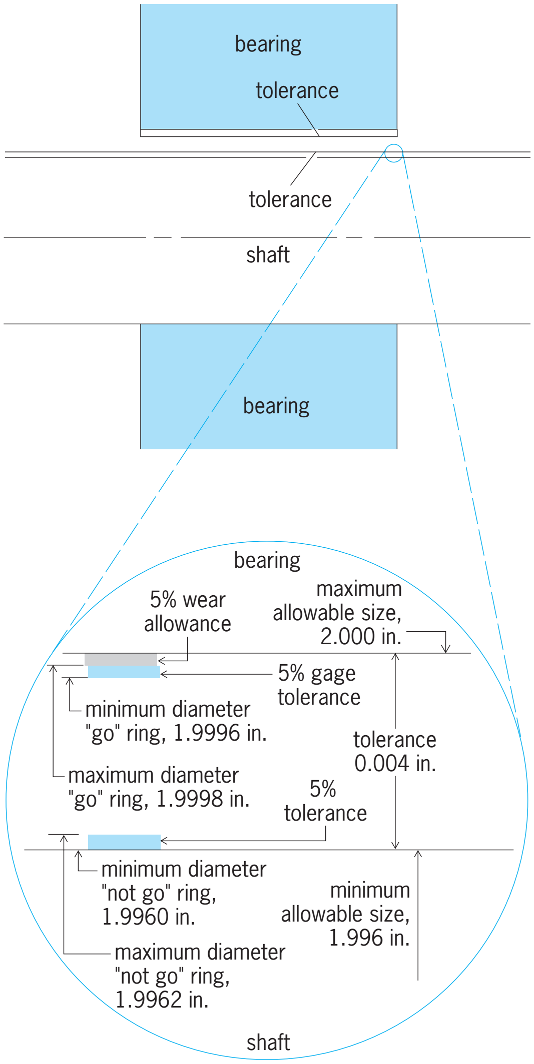


Fig. 1. Makeup of gage tolerances. 1 in. = 2.5 cm.

12.0000 in. (30.4800 cm) can be built up. Between 0.3000 in. and 4.0000 in. (0.7620 and 10.1600 cm), three to four gage groups can be built at the same time from a set to an identical length expressed in four decimal places.

Gage blocks are measured by interferometry. Their actual sizes are determined by comparison with the wavelength of red light in the color spectrum, the ultimate standard of lineal measurement being the wavelength of light. Gage blocks translate this basic lineal standard into a practical form for shop use. *See* OPTICAL FLAT.

Ring gages. Cylindrical rings of steel whose inside diameters are ground and then honed or lapped to gage tolerance are used for checking the diameter of a cylindrical object such as a shaft. The gaging surface is sometimes chrome-plated or made of carbide for wear resistance. A single gage controls one limit of tolerance of the diameter. The "go" gage will pass over the largest acceptable part but will not pass over any larger part. The "not go" gage will not pass over any part within tolerance but will pass over an undersize part. A ring gage is superior for checking the "go" dimension because it checks an infinite number of diameters simultaneously and, to the extent of its length, straightness too. Conversely, it is not preferred as a "not go" gage; an out-of-round or bent part could have many diameters undersize but still be "accepted" because the gage would "not go" if the part were bent or if even one diameter were up to the low limit. Ring gages are made to order.

Plug gage. A cylinder of steel whose diameter is ground, then honed or lapped to gage tolerance, a plug gage is used for checking the diameter of a cylindrical cavity such as a bolt hole or bearing bore. The materials, dimensions, and tolerances of plug gages are comparable to those of ring gages. A single plug gage controls one limit of tolerance of an inside diameter. Plug gages are usually used in pairs, a "go" gage and a "not go" gage. Plug gages are also made to order.

Screw-thread gage. This type of gage may be used to inspect the pitch diameter, major diameter, minor diameter, lead, straightness, and thread angle of a screw thread (**Fig. 3**). Screw-thread gages may be plug gages, ring gages, snap gages, or indicating gages. A "go" plug or ring gage inspects all these

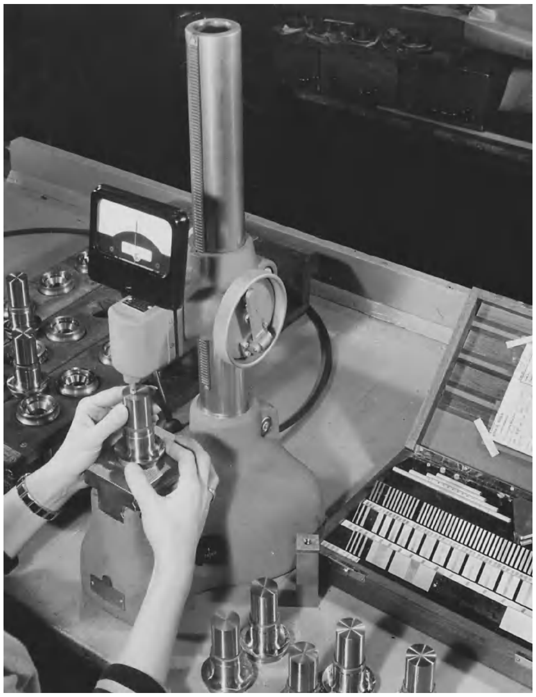


Fig. 2. General-purpose comparator of the electrical type being set with gage blocks (in box to right). (*Pratt and Whitney Co.*)

Fig. 3. Thread plug gage. (*Pratt and Whitney Co.*)

Fig. 4. Indicating gage of the electronic type being used to inspect a part. (*Federal Products Corp.*)

characteristics. A "not go" plug or ring gage checks only pitch diameter and lead. Snap or indicating gages check pitch diameter and one or more of the other characteristics. Ring gages are adjustable and can be set with a "setting" plug.

Taper pipe thread gage. Standards for pitch diameter and taper of a pipe thread were first set to control threads used in petroleum service, where safety depended on accuracy of the thread dimensions. Standards were set jointly by the American Petroleum Institute and the U.S. Bureau of Standards. Size is determined by measuring the distance that a plug gage will screw into a coupling or that a ring gage will screw onto a pipe. A manufacturer of pipe or couplings has a master plug and master ring gage certified by an approved laboratory. These are used to control working gages. As gages wear, the distance of engagement changes and limits are reset as long as thread form remains satisfactory.

Snap gage. A snap gage has two surfaces that are flat and parallel; these are spaced to control one limit of tolerance of an outside diameter or a length. A progressive snap gage has the "not go" surface just behind the "go" surface so that in a single motion both limits of tolerance can be checked. Adjustable snap gages have adjustable gaging surfaces so that the distance between gaging surfaces can be set with gage blocks and used to gage any dimension falling within the limits of the frame size.

Snap gaging is faster than ring gaging, but it checks only one diameter or length at a time. It is frequently used as a "not go" gage only, in conjunction with a ring gage for "go" limit to detect out-of-round condition in cylindrical objects. Adjustable gages have an advantage over fixed gages because allowances for wear and gage tolerance may be disregarded in their setting.

Receiving gage. A fixed gage designed to inspect a number of dimensions and also their relation to each other is termed a receiving gage because of its similarity to the cavity which receives the part in actual service. The gage checks on "go" limits. A typical example is a chamber gage which resembles the chamber of a rifle but which is made to the largest limits of size of a cartridge. Thus any cartridge which fits the receiving gage will fit any rifle chamber of that caliber.

Indicating gage. An indicating-type gage has contact points that move as they contact the part being inspected (**Fig. 4**). The movement is amplified on an indicator whose scale designates the limits of tolerance. Initial setting is done with a master or check gage. An indicator often may be substituted for a fixed or adjustable contact in which a single dimension is being checked. Snap gages are commercially available with indicators instead of fixed contact points. Indicators are also used for inspecting out-of-roundness, centerline runout, or taper, by rotating the part on one axis or diameter and indicating the "run out" of another axis or diameter. Indicating gages do not require allowances for wear or gage tolerance in their setting. *See* INSPECTION AND TESTING.

Rush A. Bowman

Bibliography. C. W. Kennedy and S. D. Bond, *Inspection and Gaging*, 6th ed., 1987; A. L. Window and G. S. Holister, *Strain Gauge Technology*, 1994.

Gain

The relative change in quantity, such as signal voltage or power, between two locations in a linear system, where the quantity at the second location depends on the quantity at the first location. It is usually expressed as the ratio of the magnitude of an output quantity to the magnitude of the associated input quantity. Equation (1) expresses the gain G as

$$G = \frac{|X_{\text{out}}|}{|X_{\text{in}}|} \qquad (1)$$

illustrated in **Fig. 1**, where $|X_{\text{out}}|$ and $|X_{\text{in}}|$ denote the magnitude of the output and input signals, respectively. Since signal gain values can be very large, they are often expressed in decibels (dB), as in Eq. (2). An operational amplifier providing a voltage gain of 100,000 has a voltage gain of 100 dB. If the input and output quantities have units of power, the decibel definition is 10 $\log_{10} G$. *See* AMPLIFIER; DECIBEL; OPERATIONAL AMPLIFIER.

$$\text{Signal gain in decibels} = 20 \log_{10} G \qquad (2)$$

The gain function that relates the input and output quantities often depends upon the frequency of the input excitation. Bode diagrams are commonly used to represent the response of a linear system to sinusoids with varying frequency. For a given sinusoidal input signal to a linear system, the output differs in

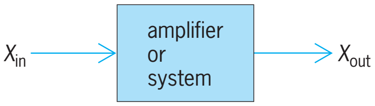


Fig. 1. Gain by an amplifier or other system.

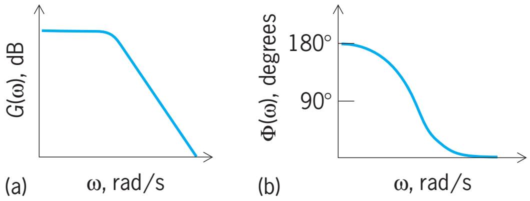


Fig. 2. Bode diagrams. (*a*) Gain plot. (*b*) Phase plot.

amplitude and phase (shift). Assuming the system in Fig. 1 is linear, if Eq. (3) holds, then Eq. (4)

$$X_{\text{in}} = A\cos(\omega t) \qquad (3)$$

$$X_{\text{out}} = AG(\omega)\cos[\omega t + \Phi(\omega)] \qquad (4)$$

follows, where ω is the radian frequency, $G(\omega)$ is the frequency-dependent gain, and $\Phi(\omega)$ is the frequency-dependent phase shift. The Bode diagrams show the gain $G(\omega)$ and the phase shift $\Phi(\omega)$ as functions of sinusoidal input frequency ω. **Figure 2** shows typical Bode gain and phase plots.

Gain functions and their graphical representations (Bode plots) are widely used in the analysis and design of mechanical and electrical systems. Besides gain and phase shift, information on system stability and other features is readily obtained from Bode plots. Kirk D. Peterson

Bibliography. T. G. Beckwith, J. H. Lienhard, and R. D. Marangoni, *Mechanical Measurements*, 5th ed., 1993; B. C. Kuo, *Automatic Control Systems*, 7th ed., 1995; M. E. Van Valkenburg, *Network Analysis*, 3d ed., 1974.

Galaxy, external

One of the large self-gravitating aggregates of stars, gas, and dust that contain a large amount of the visible baryonic matter in the universe. Typical large galaxies have symmetric and regular forms, are about 50,000 light-years (3×10^{17} mi or 5×10^{17} km) in diameter, and are roughly 3×10^{10} times more luminous than the Sun. The stars and other material within a galaxy move through it, often in regular rotation, with periods of a few hundred million years. The characteristic mass associated with a large galaxy is a few times 10^{12} solar masses. (The solar mass is 4.4×10^{30} lb or 2×10^{30} kg.) Galaxies often occur in groups or clusters containing from a few to many thousands of individual galaxies and ranging in size from a few hundred thousands to tens of millions of light-years. The nearest galaxy to the Milky Way Galaxy, the Sagittarius Dwarf Galaxy, is about 80,000 light-years (4×10^{17} mi or 6×10^{17} km) away; the farthest, more than 1×10^{10} light-years (6×10^{22} mi or 1×10^{23} km). Galaxies are the landmarks by which cosmologists survey the large-scale structure of the universe. *See* BARYON.

Composition. Most of the hundreds of billions of stars making up a galaxy are not generally individually observable because they are too faint and distant. Only the brightest stars in the nearest galaxies can be observed directly with large telescopes. Such stars include young, blue, and massive hydrogen-burning (that is, their energy comes from the nuclear fusion of hydrogen into helium) main-sequence stars; red, helium-burning giant stars near the end of their life cycle; evolved, luminous variable stars; and exploding (nova and supernova) stars. Although such stars may contribute most of the total visible light from a galaxy, they are few in number; most of the stars are in the form of much fainter and lower-individual-mass hydrogen-burning main-sequence stars and in the faint burned-out remnants of dead stars. The faint stars in galaxies can be studied only indirectly through the properties of their combined light. Furthermore, most of a galaxy's mass is in the form of "dark matter," a component of unknown nature that is detectable only through the effects of its gravitational pull on the stars of a galaxy. *See* DARK MATTER; NOVA; STELLAR EVOLUTION; SUPERNOVA.

The stars in a galaxy include many mixes of types of stars, depending on the evolutionary history of the galaxy. In the Milky Way Galaxy there are two general types of stellar populations: One type (population I) is characterized by the presence of young stars and by ongoing star formation. It is usually associated with the presence of gas. The second type (population II) shows an absence of gas and young stars as well as other indications that star formation ceased long ago. The Sun is a population I star. *See* STAR; STELLAR POPULATION.

Galaxies contain gas (mostly un-ionized hydrogen) in amounts varying from essentially zero up to a considerable fraction of their total mass. Dust in galaxies, although small in mass (typically 1% of the gas mass), is often dramatic in appearance because it obscures the starlight. *See* INTERSTELLAR MATTER.

Galaxies the size of the Milky Way Galaxy and larger have been confirmed to be dominated by dark matter, a form of matter whose nature is still unclear and whose existence has been confirmed only by gravitational effects on the surrounding visible matter. The dark matter distribution in galaxies increases radially toward the center of the galaxy and falls off steeply toward the outskirts of the galaxy. The existence of dark matter in galaxies smaller than the Milky Way Galaxy, so-called dwarf galaxies, has also been confirmed in several cases, especially the nearby low-surface-brightness dwarf galaxies in the constellations Leo and Draco.

Form and size. Galaxies generally display strikingly regular forms. The most common form is a disk with a central bulge. The disk is typically 100,000 light-years (6×10^{17} mi or 1×10^{18} km) in diameter and only about 1000 light-years (6×10^{15} mi or 1×10^{16} km) thick. Its appearance is characterized by radially decreasing brightness with a superposed spiral or bar pattern, or both (**Figs. 1** and **2**). The central bulge may vary in size from hundreds to many thousands of light-years. Such galaxies are classified as spirals (S) and subclassified a, b, or c (for example, Sa) to distinguish increasingly open spiral structure and small bulge size. The disks of these galaxies are dominated by population I stars, while their bulges contain mainly population II stars. The Milky Way Galaxy is an Sb type. *See* MILKY WAY GALAXY.

Many spiral and irregular galaxies have a nearly linear feature in their central regions. Called barred galaxies, these objects can otherwise fit into the general scheme of galaxy types. The letter B is added after the S in the classification of spiral galaxies that contain conspicuous barlike features (for example, M61 is an SBc galaxy). It is likely that almost all spiral galaxies have at least some stars moving in barlike orbits.

Another common type of galaxy is an ellipsoid with radially decreasing brightness. These galaxies are classified as ellipticals (E) and subclassified according to their axial ratios by a number from 0 (E0 = round; **Fig. 3**) to 7 (E7 = 3-to-1 axial ratio). Additionally, ellipticals are subclassified according to the shape of their lines of constant brightness on images, so-called isophotes. Very luminous E galaxies show isophotes with boxlike shapes (boxy E) while less luminous E galaxies show isophotes with disklike shapes (disky E). Ellipticals may vary in size from thousands to several hundred thousand light-years. They are most commonly found in clusters of galaxies and often contain a hot gas halo and in about half of all cases appreciable amounts of dust. The brightest galaxies are usually ellipticals. They are dominated by population II stars.

Other, rarer forms of galaxies include a transition class called S0 that has a disk superimposed on an otherwise elliptical type of light distribution, and an irregular (Irr) class composed of galaxies with chaotic forms (Fig. 2) and generally low total luminosity.

Exotic galaxy types. Some galaxies lie outside the normal range of morphologies. Most of these galaxies have suffered some disturbing event, such as a gravitational encounter, a merger, or violent nuclear activity. Other such galaxies seem to have started out with anomalous characteristics, such as the low-surface-brightness galaxies. Those with activity in their nuclei are described below.

Starbursters. One of the more spectacular examples of exotic galaxies is the starbursters, galaxies that are presently manufacturing stars at an unusually vigorous rate. It is now known that some gravitational impulse has triggered the unusual star formation activity in at least most cases. The burst is a temporary condition and the galaxies now bursting must have spent most of their lives in a more quiet condition. *See* STARBURST GALAXY.

Low-surface-brightness galaxies. Another type of exotic galaxy is the low-surface-brightness galaxies, star systems that have such a low spatial density of stars that they are almost invisible. The Sculptor dwarfs have many characteristics similar to the globular star clusters (very old stars, cluster-type variable stars, and smooth stellar distribution) but are millions of times less dense, and are dominated by dark matter. Other low-surface-brightness galaxies have a composite population of stars. For example, at least two star-forming events occurred in the Carina dwarf, one about 13×10^9 years ago and one about 7×10^9 years ago. *See* STAR CLUSTERS.

Another type of low-surface-brightness galaxy includes extreme irregular and spiral galaxies, which have some of the structural properties of the normal examples of these types but are so faint that it is difficult to detect them against the sky's brightness. A significant fraction of the mass of the universe may be in the form of these nearly invisible galaxies.

Fig. 1. Great Spiral Galaxy in Andromeda (M31, NGC 224) and its two small elliptical companions (NGC 205 and 221), photographed with the 48-in. (122-cm) Schmidt telescope. (*California Institute of Technology/Palomar Observatory*)

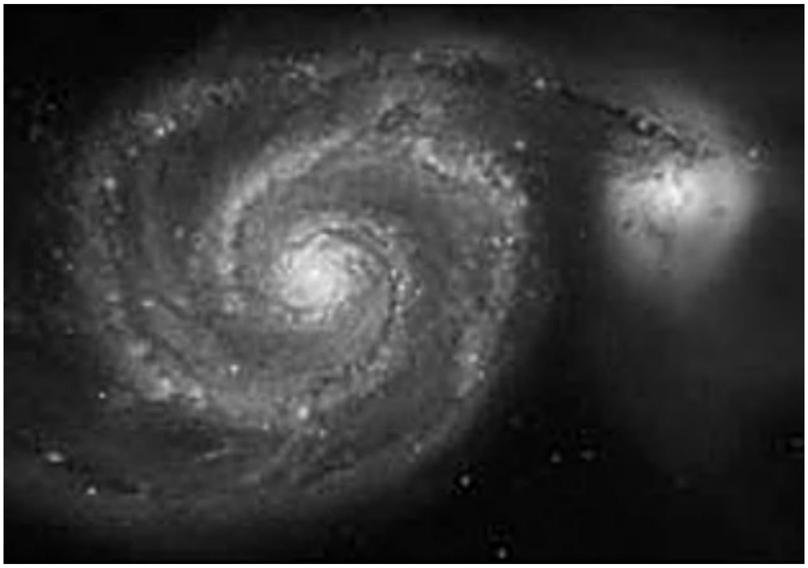


Fig. 2. “Whirlpool” Galaxy (NGC 5194), type Sc, and a companion irregular satellite (NGC 5195). [*NASA, ESA, S. Beckwith (Space Telescope Science Institute), The Hubble Heritage Team (Space Telescope Science Institute/AURA)*]

Fig. 3. E0 galaxy M87 (NGC 4486) in the Virgo cluster, which is in the constellation of Virgo. This galaxy is a source of radio emission, and it has an active nucleus. (*California Institute of Technology/Palomar Observatory*)

Internal motions. The stars and interstellar matter in a galaxy revolve around the center of mass of the galaxy, which is often a bright nucleus. In a spiral galaxy such as the Milky Way there are generally two types of stellar orbits. The stars in the flat plane tend to have nearly circular orbits, while the stars in the bulge and halo have more highly elliptical orbits. All of the motions are the result of the gravitational interactions of the stars with each other. The galaxies are supported against gravitational collapse by these motions in the same sense in which the planets of the solar system are kept from falling into the Sun by their orbital motions. Usually, the inner regions in the plane of a spiral galaxy undergo nearly solid-body rotation (velocity proportional to radius), while the outer regions rotate differentially (velocity constant). The total velocity of the material within galaxies varies from one to several hundred miles per second. Typical orbital periods for stars are several hundred million years.

The distribution of kinetic energy into randomly oriented and circular rotational motions varies with the galaxy type. The disk of a spiral galaxy may have only about 1% of its total kinetic energy in random motions, while boxy E galaxies in contrast to disky E galaxies may have most of their kinetic energy in their random component.

Luminosities. The number of galaxies with total luminosities L is roughly proportional to L^{-y}, where y is between 1 and 1.5 for luminosities less than about 3×10^{10} solar luminosities. The number is exponentially cut off for higher luminosities. The brightest observed galaxies are fainter than 2×10^{11} solar luminosities; the faintest, brighter than about 1×10^{6}. The Milky Way Galaxy's total luminosity is roughly 1×10^{10} times greater than the Sun's.

The distribution of luminosities is such that while there are very many faint galaxies, they do not contribute a large fraction of the total light given off by galaxies. Only the brighter galaxies, visible for great distances through space, can be observed easily and in great numbers. It is generally found that the masses of galaxies are roughly proportional to their luminosities.

At each luminosity the fraction of different galaxy types strongly varies with luminosity. The most luminous galaxies are mainly ellipticals and the less luminous mainly spirals. Additionally the number of elliptical galaxies at all luminosities increases in environments where the density of galaxies is high.

Clustering. Although galaxies are scattered through space in all directions for as far as they can be observed, their distribution is not uniform or random. Most galaxies are found in associations (**Fig. 4**) containing from two to hundreds of individual bright galaxies and at least 10 times as many fainter dwarf galaxies. The E and S0 galaxies tend to be concentrated in large clusters more strongly than spirals. The Milky Way Galaxy and the Andromeda Nebula (Fig. 1) are members of a cluster called the Local Group. *See* ANDROMEDA GALAXY; LOCAL GROUP.

Like the stars within a galaxy, the galaxies within a cluster move about under the influence of their mutual gravitational attraction. The motions are generally randomly oriented and the cluster shows little evidence of rotation. Typical velocities range from about a hundred up to a thousand miles per second or more.

On scales larger than individual small groups and rich clusters, the distribution of galaxies through space is still not random. This very large scale structure in the galaxy distribution is usually referred to as superclustering to indicate that it involves the higher-order clustering of the individual first-order associations of galaxies. The largest structures generally do not have the shape of spherical clusters, but rather the concentrations of galaxies tend to be filamentary in nature. Between filaments there are often large-scale voids, in which few galaxies are found.

This very large scale structure is generally attributed to the hierarchical clustering that resulted from the gravitational instability from primordial fluctuations immediately following the big bang. These would have systematically formed small galaxies first, followed by larger and larger concentrations of more massive galaxies. Evidence for the existence of the density fluctuations that are inferred to have seeded structure formation has come from the discovery and mapping of the relic temperature fluctuations in the cosmic microwave background. Indeed, computer simulations of conditions in the early universe indicate that the observed large-scale distribution of galaxies and galaxy clusters is a natural result of the evolution of the universe after the big bang, when the earliest structures began to form in what was then a much denser universe. *See* COSMIC BACKGROUND RADIATION; COSMOLOGY; UNIVERSE.

Active nuclei. In the very central regions (sizes at least as small as a light-year, 6×10^{12} mi or 1×10^{13} km) of galaxies, violent behavior is often observed. This activity is manifested in many ways, including the high-velocity outflow of gas, strong nonthermal radio emission (implying relativistic particles and magnetic fields), intense and often polarized and highly variable radiation at infrared, optical, ultraviolet, and x-ray wavelengths, and ejection of jets of relativistic material. In the most extreme cases the energy in the nuclear activity surpasses that in the rest of the galaxy combined. These phenomena are generically referred to as nuclear activity, and the objects that exhibit them are called active galactic nuclei.

One of the most prominent characteristic features of active galactic nuclei is the ejection of large masses of high-temperature gas at great velocities. Characteristic temperatures and velocities are in the range of tens to hundreds of thousands of kelvins and thousands to over 10,000 mi/s (16,000 km/s). Total gas masses exceeding 1×10^6 times that of the Sun may be involved. These powerful gas flows reveal themselves as bright and broad spectral lines.

A second characteristic feature of active galactic nuclei is the emission of radiation over a wide range of different wavelength bands, from the radio band to gamma rays. An object such as an ordinary star, which emits radiation because it is hot, does so in a characteristic wavelength band. Such radiation is called thermal, and the characteristic wavelength is determined by the object's temperature. The typical active galactic nucleus emits a quite different sort of radiation, called nonthermal radiation, implying that it is produced by a quite different mechanism. *See* HEAT RADIATION.

Another indication that the radiation is of an unusual, nonthermal origin is the dramatic brightness changes of active galactic nuclei over relatively (by astronomical standards) short periods of time. Most active galactic nuclei show moderate changes over time scales of months to years, and some show dramatic variations (by factors of 2 to 10 or more) over times ranging down to a few hours. Since sources of radiation cannot generally change their brightnesses in times much shorter than that required for light to travel across them, these variations imply that the radiation from an active galactic nucleus arises in a very small region, in extreme cases no larger than the solar system. This is a fantastically small volume, considering that the total radiative power output can rival or exceed that of an entire galaxy many tens of thousands of light-years across.

Another important and peculiar phenomenon exhibited by active galactic nuclei is the emission of highly directional jets of relativistic plasmas and magnetic fields. These jets contain elementary particles, particularly electrons, moving at velocities near the speed of light, and often extend far outside the nucleus and even beyond the whole body of the galaxy. *See* PLASMA (PHYSICS).

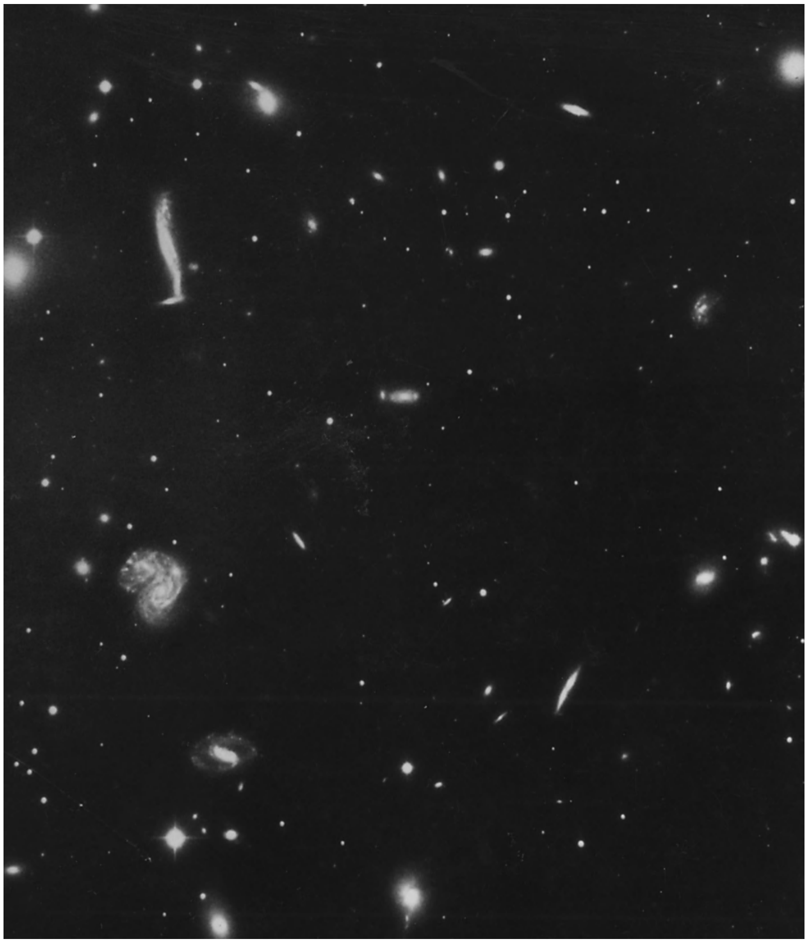


Fig. 4. Clustering characteristics of galaxies. This association of galaxies in the constellation Hercules includes many different types of galaxies. (*California Institute of Technology/Palomar Observatory*)

There are a variety of classes of active galactic nuclei. The Seyfert galaxies display the broad emission lines produced by the rapid outflow of hot gas but frequently do not exhibit much radio-wavelength emission. Another complementary class shows strong radio emission but weak or absent emission lines. Yet another class (BL Lac objects, often referred to as blazars) also shows only weak emission lines but is often extremely variable. When active galactic nuclei achieve such great luminosities that they dominate that of the rest of the galaxy, they are sometimes referred to as AGNs. Quasars are the most extreme sort of active galactic nuclei, having emission so intense that the ordinary galaxy in which they exist is entirely lost in the glare of the nuclear emission. It appears that active galactic nuclei in general, and certainly quasars, were much more common during the early history of the universe than they are at present. If so, most or all large galaxies may contain the burned-out remnant of an active nucleus in their centers, although only about 1% still show detectable activity. *See* QUASAR.

Perhaps the most intriguing question concerning active galactic nuclei is that of the nature of the energy source that drives all of their diverse phenomena. While there is no certain answer to this question, there is a widely accepted best guess or consensus model that appears consistent with all that is known about active galactic nuclei. The basic idea of this model is that active galactic nuclei are powered by the energy released when matter falls into a massive black hole occupying the center of a galaxy. These black holes are found to have masses in the rough range of 10^6–10^9 solar masses and to have formed because of the high density of material expected to accumulate at the center of a galaxy due to its gravitational field. Such a black hole will continue to accrete any gas that finds its way into the vicinity. As such gas falls toward the black hole, its angular momentum will cause it to take up a nearly circular orbit in a disk of material surrounding the black hole. This disk (called an accretion disk) will slowly inject gas into the black hole. As the gas approaches the black hole, the latter's enormous gravitational field will compress and heat the gas to very high temperatures, causing it to radiate. Intense jet radio emission, powered by energy released during infall onto the central black hole, is ejected along the minor axis of the accretion disk. Depending on the viewing angle of the observer, the resulting morphology can account for a wide variety of active galactic nuclei, radio galaxies, and quasars. A given mass of gas can release 10 or more times as much energy in this way as it could if it were used as nuclear fuel in a star or a reactor. A gas infall rate onto the central engine of several solar masses of gas per year suffices to power the most luminous active galactic nuclei in the universe. There is ample gas available in the interstellar medium to act as a fuel supply. Gas may be driven into the central regions of the galaxy, perhaps following a merger, where it can supply and activate the nucleus. The mechanisms that covert the thermal radiation generated in this way into the nonthermal radiation and relativistic plasmas observed in active galactic nuclei are not well known. *See* ASTROPHYSICS, HIGH-ENERGY; BLACK HOLE.

Paul Hodge; Edwin L. Turner; Joyce B. Turner

Correlations. A remarkable number of well-established correlations exist between structural properties of elliptical and spiral galaxies that are indicative of their origin. As mentioned above, spiral galaxies rotate like a solid body in their inner parts and differentially in their outer parts. A maximum in the rotation velocity usually occurs at the transition point between these two regimes. This maximum rotation velocity is strongly correlated to the luminosity of the spiral galaxy in such a way that it increases with luminosity. A similar correlation has been found for elliptical galaxies between the amount of random motion (the so-called velocity dispersion) of their stars and their luminosity. More luminous ellipticals show larger random motion of their stars.

The integrated color of the stars in spirals and ellipticals is redder in the center and becomes bluer going radially away from the center. The reason for this behavior is twofold. Stars that grow old become redder and stars that are enriched with elements more heavy than helium tend to be redder, also. The interpretation of the data therefore suggests that the stars in the center are older and that the percentage of elements heavier than helium increases toward the center. This latter influence is verified by the strength of spectroscopic features in the galactic starlight that measure the abundance of iron and other heavy elements.

Additionally the global percentage of heavy elements in a galaxy of any type increases with its luminosity. This is explained by the larger gravitational potential generated by the larger amount of mass in more luminous galaxies. This potential is very effective in keeping the heavy elements produced during the life cycles of stars that subsequently are ejected into the interstellar medium by supernovae. The heavy elements are subsequently incorporated in the next generation of new stars and cause them to be redder. This and an old generation of stars are responsible for the observed correlation that more luminous galaxies are redder.

Elliptical galaxies over a wide range of luminosities are compared according to their central surface brightness μ_0, central velocity dispersion σ_0, and effective radius R_0, which is the radius from within which half the light of the galaxy is emitted. The comparison shows that they do not occur in all possible combinations but occupy a two-dimensional plane, called the fundamental plane, in this three-dimensional parameter space. Thus, only certain combinations of these three parameters are realized by nature. This correlation is usually attributed to the condition that the galaxies are in an energetically relaxed state.

It is widely believed that bulges of both spirals and elliptical galaxies harbor supermassive black holes with masses up to several times 10^9 solar masses at their centers. A striking correlation between the central velocity dispersion of ellipticals and bulges of spirals with the central black hole mass has been established, which hints towards a connection in the formation of bulges and supermassive black holes. Spheroids with larger velocity dispersion have more massive black holes.

Sadegh Khochfar; Joseph Silk

Bibliography. J. Binney and M. Merrifield, *Galactic Astronomy*, Princeton University Press, 1998; J. Binney and S. Tremaine, *Galactic Dynamics*, 1987; D. M. Elmegreen, *Galaxies and Galactic Structure*, Prentice Hall, 1998; J. H. Krolik, *Active Galactic Nuclei: From the Central Black Hole to the Galactic Environment*, Princeton University Press, 1999; A. Sandage and J. Bedke, *The Carnegie Atlas of Galaxies*, 2 vols., 1994; A. Sandage, *The Hubble Atlas of Galaxies*, Carnegie Institution, 1961; S. van den Bergh, *Galaxy Morphology and Classification*, Cambridge University Press, 1998; S. van den Bergh, *The Galaxies of the Local Group*, Cambridge University Press, 2000; W. H. Waller and P. W. Hodge, *Galaxies and the Cosmic Frontier*, Harvard University Press, 2003.

Galaxy formation and evolution

How the diverse array of galaxies that are now observed originated and evolved into their present form is a topic of intense speculation.

Evidence from structural properties. Some clues can be discerned in certain structural properties of galaxies. The most plausible explanation for the smooth and round light distribution of an elliptical galaxy is that the stars formed out of a collapsing gas cloud. The rapidly changing gravitational pull experienced by different stars as the collapse proceeds has been shown by computer simulations to rearrange the stars into the observed shape.

The highly flattened disks of spiral galaxies must have formed during a similar collapse, but it is believed that most star formation did not occur until the rotating gas cloud had already flattened into a pancakelike shape. Had the stars formed at an earlier stage of the collapse, their rapid motions would have led to the formation of an elliptical galaxy. However, if the cloud stays gaseous until it flattens, much of the kinetic energy of the collapse is radiated away by gas atoms. Subsequent star formation is found to maintain a highly flattened, disklike shape, characteristic of spiral galaxies.

The flattening occurs in part because of the centrifugal forces in the rotating cloud. A confirmation of this picture has come from the discovery that elliptical galaxies rotate much less rapidly than spiral galaxies. This raises the question of the origin of the rotation itself. A natural explanation seems to lie in the action of the gravitational torques exerted by neighboring protogalaxies.

Evidence from composition. Another aspect of galaxies that has evolutionary significance is their composition, and, in particular, the distribution of heavy elements. The amount of heavy-element enrichment can be inferred from the color of the starlight, blue stars being metal-poor. Galaxies are found to be significantly bluer in their outermost regions and redder toward their central nuclei. The explanation seems to be that galaxies formed out of collapsing gas clouds that formed stars in a piecemeal fashion. As stars formed, they evolved, underwent nuclear reactions, produced heavy elements, and eventually shed enriched material (some stars even exploding as supernovae). Successive generations of stars formed out of the debris of earlier stars, and in this way the stellar content of galaxies systematically became enriched. The greatest enrichment would naturally occur toward the center of a galaxy, where the gaseous stellar debris tended to collect. *See* NUCLEOSYNTHESIS.

Interactions. Observations reveal many systems of interacting galaxies and close pairs of galaxies, which are possible candidates for later interactions. As galaxies move about within clusters, they will occasionally pass very near one another or even collide directly. Possible outcomes include loss of material from the galaxies' outer regions, transfer of material from one galaxy to another, merger of the two galaxies, modification of the galaxies' forms by tidal perturbations, and loss of gas and dust due to collisional heating. Whether a merger occurs mainly depends on the relative velocity difference of the two galaxies. If they pass each other too fast, the gravitational drag between them will not be efficient enough to change their trajectory and the passage does not result in merging.

Observations show an increase of interactions in the last few 10^9 years. The random velocity of a galaxy inside an association of galaxies increases with the mass connected to that association. Once the association forms and the galaxies within it start moving under its gravitational influence, the relative velocities in encounters of galaxies will become too high for mergers to occur and the number of mergers decreases. Looking back in time, the increase of interactions is stronger in clusters of galaxies than in environments with few galaxies. This is a consequence of the higher number density of galaxies, which increases the probability of having an encounter.

Galaxies undergoing mergers experience dramatic morphological changes. Due to tidal forces the merging galaxies start deforming and develop very prominent features, most notably the so-called tidal arms. Mergers can trigger periods of intense star formation. Gas which was available in the disks of the progenitor spiral galaxies will be driven to the centers of the remnant galaxies and start forming stars in a starburst. It has been proposed that ultraluminous infrared galaxies like the Antennae galaxies (**Figs. 1** and **2**) are just galaxies undergoing mergers in which an extensive starburst occurs. Theoretical considerations suggest that the outcome of the interaction

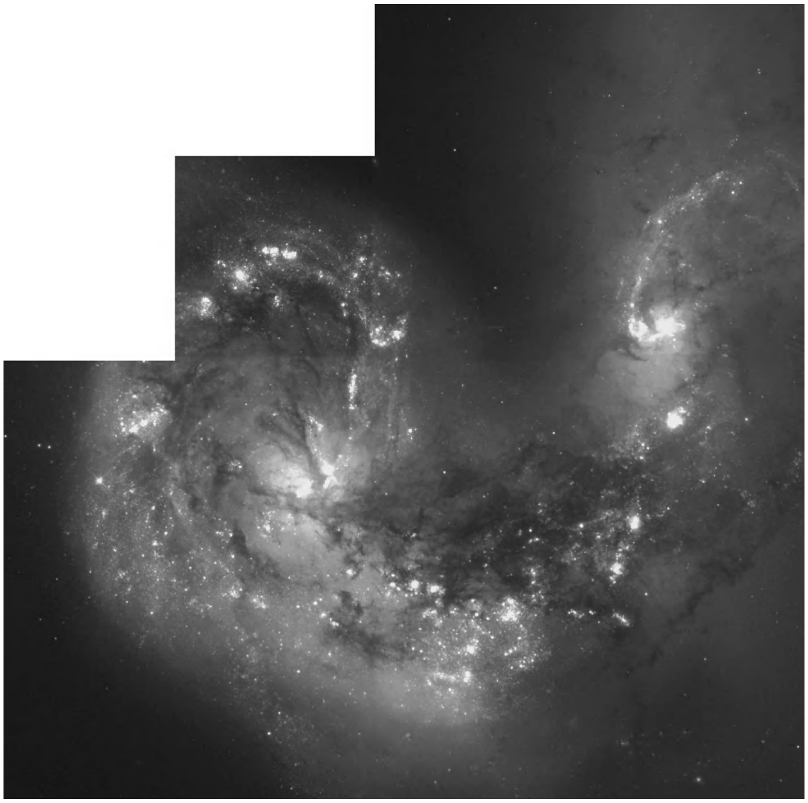


Fig. 1. Centers of the Antennae Galaxies (NGC 4038 and NGC 4039) seen in visible light. This pair of interacting (colliding) galaxies shows many sights of prominent star formation. [*B. C. Whitmore Space Telescope Science Institute), NASA*]

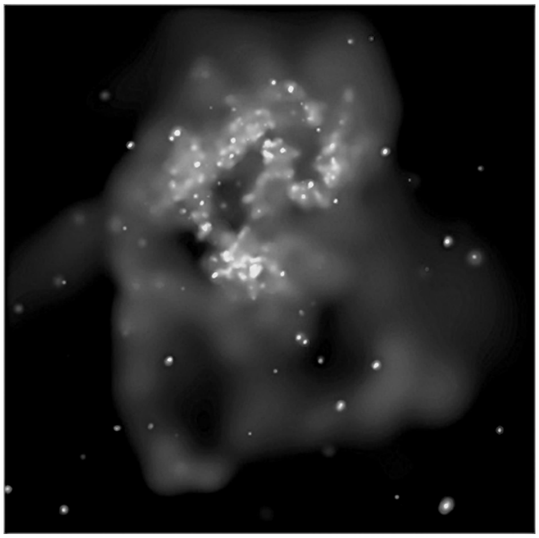


Fig. 2. Antennae Galaxies seen in x-rays. Many bright spots are seen which could be connected to black holes. A very prominent halo of hot gas is also seen. (*NASA, Chandra X-ray Observatory; Smithsonian Astrophysical Observatory; G. Fabbiano et al.*)

between two galaxies of similar size will be an elliptical galaxy. This is one of the most favored formation mechanisms for elliptical galaxies. *See* STARBURST GALAXY.

Origin. An outstanding and unresolved issue concerns the origin of the primordial gas clouds out of which the galaxies evolved. The cosmic background radiation yields a glimpse of the universe prior to the epoch of galaxy formation. The universe is now completely transparent to this radiation. However, about 500,000 years after the big bang, the radiation was sufficiently hot that matter was ionized, and the matter was also sufficiently dense to render the universe completely opaque to the radiation. To observe the background radiation now (some 10^{10} years later) is to see back to this early epoch, known as the decoupling epoch: at earlier times, matter and radiation were intimately linked, and subsequently the radiation propagated freely until the present time. *See* BIG BANG THEORY; COSMIC BACKGROUND RADIATION.

Theory of formation from fluctuations. The cosmic background radiation is very uniform, but fluctuations were discovered in 1992 by the *Cosmic Background Explorer (COBE)* satellite. These are at a level of 1 part in 10^5, and are on angular scales of several degrees. The precursor fluctuations of the primordial inhomogeneities that gave rise to the observed structures in the universe would be on scales of degrees, for the largest superclusters and voids, to arcminutes, for galaxy clusters. A definitive measurement is not yet available for these angular scales, but without such primordial fluctuations galaxies could not have formed. The mutual action of gravity exerted between these infinitesimal fluctuations results in their gradual enhancement. Eventually, great gas clouds develop that will collapse to form galaxies. The required amplitude for these primordial seed fluctuations must be of the order of 1 part in 10^5, precisely what is measured on larger scales, to within uncertainties of at most a factor of 2.

Numerical simulations of galaxy clustering have enabled the spectrum of fluctuation length scales and amplitudes to be inferred. Galaxies are not randomly distributed, as would be "white" noise; rather, they are correlated. Given a galaxy at an arbitrary position, at a distance away, there is an excess probability, above random, of finding another galaxy. These correlations are large on scales less than 5 megaparsecs (1.0×10^{20} mi or 1.5×10^{20} km), and are measured out to 20 Mpc (4×10^{20} mi or 6×10^{20} km). The parent fluctuations that gave rise to the galaxies must be similarly correlated, although the amplitude of the effect was much less.

A theory of the very early universe, first proposed in 1980, provides an explanation of the distribution of amplitudes of the fluctuations with scale, but accounts only qualitatively for their strength. According to this theory, the initial stages of the big bang were characterized by a period of rapid inflation during the first 10^{-35} s of the expansion. One consequence of an inflationary epoch is that quantum-statistical fluctuations are amplified up to scales of galaxies and of clusters of galaxies. The predicted distribution of fluctuations is initially the same, on all scales, from that of galaxy clusters to the observable universe. On the largest scales, where there has been little time to develop deviation from the initial conditions, the distribution of cosmic microwave background fluctuations measured by *COBE* in 1992 and in many subsequent experiments, especially the *Wilkinson Microwave Anisotropy Probe (WMAP)*, is approximately consistent with that predicted by the simplest inflationary cosmological model. Thus, the most simple of cosmologies may contain the nascent seeds of future galaxies. *See* INFLATIONARY UNIVERSE COSMOLOGY; WILKINSON MICROWAVE ANISOTROPY PROBE.

Isothermal and adiabatic fluctuations. The possible fluctuations in the early universe can be categorized into distinct varieties. Of particular importance for galaxy formation are density fluctuations that are found to generally be a combination of two basic types: adiabatic and isothermal.

Primordial adiabatic fluctuations are analogous to a compression of both matter and radiation. They are generic to almost all models of the early universe. In the absence of weakly interacting dark matter, the diffusive tendency of the radiation tends to smooth out the smaller adiabatic fluctuations. This process remains effective until the decoupling epoch, and galaxy formation occurs only relatively recently. However, the dominant presence of dark matter that does not interact with the radiation other than by gravity allows fluctuations to survive on all scales in the weakly interacting dark matter. The theory of fluctuation origin does not specify the strength of the fluctuations. However, the observations of the cosmic microwave background demonstrate that some 300,000 years after the big bang when the radiation was last scattered by the matter, the amplitude of the density fluctuations amounted to only a few parts in 10^4 on galaxy cluster scales. This means that massive galaxies and galaxy clusters formed relatively recently, although small galaxies could have formed when the universe was just a tenth of its present size. Galaxies form when the gaseous matter cools and condenses in the gravity field of the dark matter, forming gas clouds that subsequently fragment into stars. *See* DARK MATTER.

In some variations of the standard model for structure formation, primordial isothermal fluctuations were also present in the very early universe. These consist of variations in the matter density, without any corresponding enhancement in the radiation density. Consequently, in the radiation-dominated early phase of the big bang, isothermal fluctuations neither grow nor decay, as the uniform radiation field prevents any motion. Once the universe becomes transparent, the matter fluctuations respond freely to gravity and grow if they are above a certain critical size. The smallest isothermal fluctuations that can become enhanced and form gas clouds contain about 10^6 solar masses. Galaxy formation occurs very early in this case.

Role of dark matter. The presence of weakly interacting dark matter is almost universally accepted by astronomers in order to account for the rotation curves of galaxies. The baryonic component of matter in the universe is known from calculations of the abundances of the light elements to amount to about 3% of the critical density for closing the universe. There is at least 10 times as much dark matter, which constitutes at least 90% of the mass of the universe. Consequently, dark matter dominates the growth of the primordial density fluctuations. The gravitational influence of this dark matter greatly aids this growth. Baryon fluctuation growth is suppressed by interactions with the radiation prior to the epoch of decoupling of matter and radiation, whereas weakly interacting particles are able to cluster freely as long as the dominant form of density is ordinary matter rather than radiation. Since the density in the very early universe was dominated by radiation, fluctuation growth in the presence of dark matter is enhanced by about a factor of 10, equivalent to the expansion factor between the epochs of ordinary matter dominance when fluctuation growth first commences and the last scattering of the radiation. The associated fluctuations in the cosmic microwave background required in order to form structures by a given epoch are reduced by a corresponding factor. The detection of cosmic microwave background temperature fluctuations at a level of about one part in 10^5, initially by the *COBE* satellite and subsequently by more than 20 experiments, means that the precursor fluctuations of the largest structures, such as galaxy clusters, have been identified, in a statistical sense, in the sky. Dark matter plays an essential role in reconciling the level of the observed fluctuations with the limited growth period available since the universe was first matter-dominated, approximately 10,000 years after the big bang. This matter consists of massive weakly interacting particles whose existence is predicted by the theory of supersymmetry. *See* BARYON; SUPERSYMMETRY; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

The implications of particle dark matter for structure formation are considerable. If the particles are massive, they are slowly moving at the onset of fluctuation growth, when the universe is first matter-dominated. Such particles are called cold dark matter. As structure develops, cold dark matter clusters, first on the smallest scales, then on progressively larger scales. This leads to a bottom-up scenario of hierarchical clustering. Unique predictions are made for both the microwave background fluctuations and the density fluctuations that are measured in large-scale structure studies.

Reconciliation of large-scale structure and galaxy formation with cold dark matter has proven remarkably successful on the largest scales. Most data point to a universe in which the density of cold dark matter is about 30% of the critical value. The cosmic microwave background temperature fluctuations demonstrate that the universe is at critical temperature to account for the locations of the observed angular peaks in the fluctuation distribution observed on the microwave sky. In a universe that has a near-Euclidean geometry, the positions of these peaks are displaced because of the bending of the light rays that have traversed the universe since the epoch when the microwave background photons were last scattered by the matter. A Euclidean geometry requires that the universe must be at critical density, if most of the energy density is in the form of the vacuum energy that is associated with the cosmological constant term introduced by Albert Einstein. One prediction of such a cosmological model is that the expansion of the universe is currently accelerating as a consequence of the nature of the additional energy. Data from use of distant supernovae to measure the deceleration of the universe suggest that the universe is indeed accelerating. Large-scale structure still provides a severe constraint on the nature of the dark matter. *See* ACCELERATING UNIVERSE; COSMOLOGICAL CONSTANT; DARK ENERGY.

In a universe with a critical density of dark matter, excessively strong clustering of galaxies occurs. Dark energy, which is uniform and smooth, does not participate in gravitational clustering. Three independent observational results contribute to make a strong case for a standard model of the modern universe. These are the cosmic microwave background temperature fluctuation peaks, the acceleration of the universe as inferred from the distances to remote supernovae, and the large-scale structure of the galaxy distribution. The standard model of the universe consists of 30% dark matter, 65% dark energy, and 5% baryons. Variations in the standard model introduce a component of hot dark matter. This matter consists of neutrinos that are assumed to have a small mass, sufficient to account for about one-quarter of the critical density. However, the resulting mixture of hot and cold dark matter gives poor agreement with the astrophysical data on fluctuations at all scales. *See* NEUTRINO.

According to a much less accepted alternative viewpoint, the dark matter is entirely baryonic. It consists of very low mass stars or of burnt-out stars such as white dwarfs. In this case, the matter density of the universe is only about one-tenth of the critical density, with the rest of the critical density being made up of dark energy. One then finds that adiabatic fluctuations in a baryon-dominated universe, supplemented by a subdominant admixture of hot dark matter, can result in temperature fluctuations that agree with the observational constraints. A further possible advantage of this interpretation is that by reducing the Hubble constant to about two-thirds of the currently preferred value, it is also possible to dispense with dark energy if one ignores the evidence for acceleration of the universe inferred from distant supernovae. The universe would then contain a critical density of baryons, along with some hot dark matter. Another scenario appeals to warm dark matter, for which a massive neutrino is the expected candidate. However, this option requires an admixture of isothermal fluctuations in order to allow early structure formation. The isothermal fluctuations are consistent with the cosmic microwave background data, provided they are subdominant. In this case, a complex baryon genesis scenario is required, in which matter is created with spatial inhomogeneities in the number of baryons relative to the number of photons. By far the simplest model is one in which the observed structure is seeded by primordial adiabatic density fluctuations generated during inflation. *See* COSMOLOGY; UNIVERSE.

Joseph Silk; Sadegh Khochfar

Bibliography. J. Binney and S. Tremaine, *Galactic Dynamics*, 1987; P. Coles and F. Lucchin, *Cosmology: The Origin and Evolution of Cosmic Structure*, 2d ed., 2002; S. Dodelson, *Modern Cosmology*, 2003; M. S. Longair, *Galaxy Formation*, Springer-Verlag, 1998; T. Padmanabhan, *Theoretical Astrophysics*, vol. 3: *Galaxies and Cosmology*, 2002; P. J. Peebles, *Principles of Physical Cosmology*, 1993; J. Silk, *On the Shores of the Unknown: A Short History of the Universe*, 2005; J. Silk, *The Big Bang*, 3d ed., 2000.

Galilean transformations

The family of mathematical transformations used in Newtonian mechanics to relate the space and time variables of uniformly moving (inertial) reference systems.

In the simple case of two similarly oriented cartesian reference frames, moving along their common (x, x') axis, the transformation equations can be put in the form of Eqs. (1), where x, y, z and x', y', z'

$$x' = x - vt \qquad y' = y \qquad z' = zt' = t \qquad (1)$$

are the space coordinates of a given particle, and v is the speed of one system relative to the other.

The transformation equations for cartesian reference frames having arbitrary displacements and orientations take the more general form $(x_1 = x, x_2 = y, x_3 = z)$ of Eq. (2), where a_1, a_2, a_3, a_4 and υ_1, υ_2, υ_3

$$x'_j = \sum_{k=1}^{3} c_{jk}\,(x_k - a_k - \upsilon_k t) \qquad t' = t - a_4 \qquad (2)$$

are arbitrary real numbers and the coefficients (c_{jk}) are constants. The matrix $C = [c_{jk}]$ is real and orthogonal, so that it satisfies the condition $C^{-1} = C_t$, where C^{-1} and C_t are the inverse and transposed matrices of C.

The galilean transformations form a 10-parameter group which can be generated from translations of the space and time coordinates, rotations of the space coordinate frame, and transformations to moving reference frames. *See* FRAME OF REFERENCE; SPACE-TIME.

E. L. Hill

Gallbladder

A hollow muscular organ, present in humans and most vertebrates, which receives dilute bile from the liver, concentrates it, and discharges it into the duodenum. It also participates in the enterohepatic (re)circulation of bile, and in secretion and removal of conjugated xenobiotics, including radiopaque substances taken orally or intravenously. Although not a vital organ, it stores bile, regulates biliary tract pressures, and, when diseased, enhances precipitation of various constituents of the bile as gallstones.

Extrahepatic biliary tract. The system of bile ducts lying outside the liver is known as the extrahepatic biliary tract.

In humans (**Fig. 1**) right and left hepatic ducts empty into the common hepatic duct, which

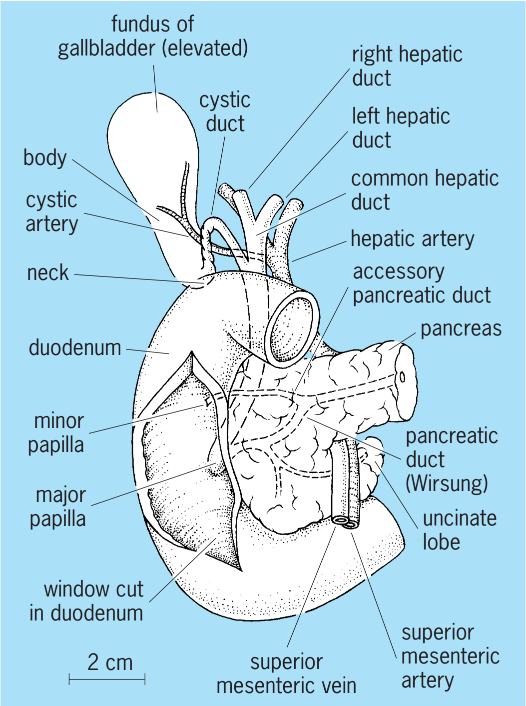


Fig. 1. Extrahepatic biliary tract in humans.

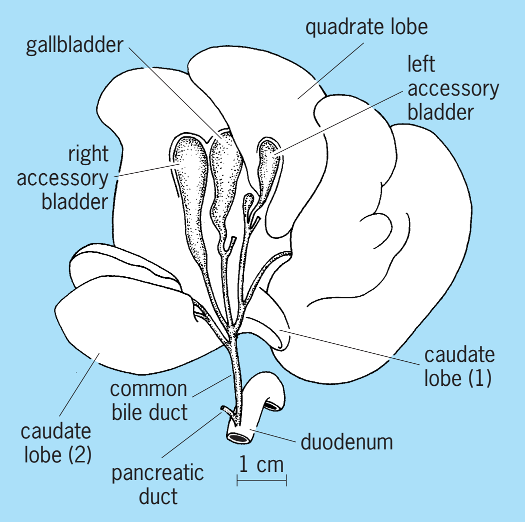


Fig. 2. Liver and accessory gallbladders of the ductule type in the cat. (*After E. A. Boyden, Amer. J. Anat., 38:177–231, 1926*)

continues to the duodenum as the common bile duct, or ductus choledochus. The gallbladder and cystic duct thus appear to be accessory organs and therefore are removable. However, they are converted into main-line structures by the presence of a sphincter at the choledochoduodenal junction. Tonic contraction of this sphincter between meals forces the bile to back up into the gallbladder. *See* LIVER; PANCREAS.

In most other vertebrates (**Fig. 2**) essentially similar relations exist except when the gallbladder is absent, but there is considerable variation in proportion and arrangement of ducts, including the pancreatic ducts. *See* PANCREAS.

Absence of gallbladder. In certain species, and as a rare anomaly in other, the gallbladder may be absent. The organ first appears in cyclostomes. It is present in all fishes, in amphibians, in reptiles, and in most birds. It is completely absent in ostriches and several other species of bird families which usually have a gallbladder (for example, it is absent in New World parrots but present in cockatoos as well as in other Australian parrots). Primitive mammals and their derivative groups—the insectivores, bats, flying lemurs, carnivores, and primates—have a gallbladder. It is absent in many later-originating herbivorous types, such as the horse, tapir, rhinoceros, elephant, peccary, deer, and camel. It is also absent in dolphins and whales. Absence of the gallbladder in the giraffe and hippopotamus is not uncommon. Conditions are most chaotic in the Eocene-developed group of rodents; among these it is missing in the common rats, castor, kangaroo rat, some porcupines, and some gophers. In humans, the gallbladder is usually embedded in the liver; it can, however, float free in approximately 4% of the normal population. Another anomaly, a folded fundus, can be found with an incidence of 2%. Additionally, in humans, agenesis (incomplete and imperfect development) and atresia (closure of normal openings) of the gallbladder occur, as well as its duplication. The cystic duct may be absent or duplicated. Moreover, the cystic duct may be abnormally long and spiral-shaped, joining the common bile duct anteriorly or posteriorly, close to its termination, or it may connect and enter the common bile duct close to its termination, or it may join the common hepatic duct; sometimes an accessory hepatic duct enters the gallbladder itself. The common bile duct can be duplicated or may develop accessory ducts. While ectopic gallbladders can be found associated with other anomalies and diseases, four different accessory gallbladders also occur in mammals, tending to be characteristic for a given species, as those of the human, cow, sheep, pig, and cat (Fig. 2).

Aberrant but vestigial bile ducts (of Luschka), originating in the hepatic tissues but not connected with the gallbladder lumen, can sometimes be observed in the adventitial layer.

Development of extrahepatic biliary tract. The way in which this system develops in humans is responsible for certain untoward clinical episodes. Both liver and ventral pancreas arise as outgrowths of a diverticulum of the embryonic gut (**Fig. 3**). Because the ventral pancreas is carried out on the bile duct, the latter becomes embedded in the head of the pancreas. Cancer of this gland can therefore obstruct the flow of bile, causing it to back up into the capillaries of the liver. Another consequence is the creation of a terminal chamber, the ampulla of Vater, into which bile and pancreatic ducts empty. Stones acting as a ball valve may lodge there; or spasms of the sphincter ampullae may result in reflux of bile into the pancreatic duct, a factor in pancreatitis. The left lobe of

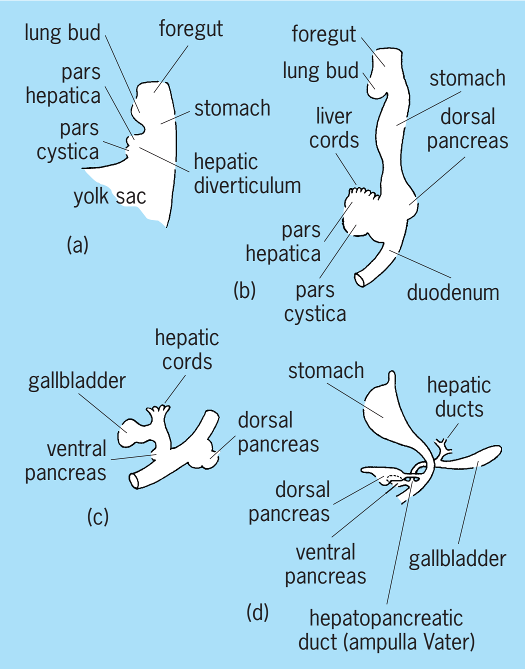


Fig. 3. Development of gallbladder (pars cystica) in human embryos. (*a*) 26 days after ovulation (horizon XII). (*b*) 28 days (XIII); 5-mm embryo. (*c*) 32 days (XV); 7-8 mm. (*d*) 35 days (XVII); 13-mm. Roman numerals indicate developmental horizons of Streeter.

the ventral pancreas may encircle the gut (annular pancreas) with consequent obstruction of the duodenum. In cats such aberrant lobes may grow up alongside the gallbladder, giving rise to a pancreatic bladder. Because of the oblique passage of the ducts through the intestinal wall, herniations or duodenal diverticula of the mucous membrane may develop. *See* LIVER DISORDERS; PANCREAS DISORDERS.

Choledochoduodenal junction. The oblique pathway through the duodenal wall traversed by the common bile duct and pancreatic duct is the choledochoduodenal junction. In humans this is surrounded by a musculus proprius which is subject to the same hormonal control as the gallbladder. This sphincter of Oddi, developing from embryonic connective tissue at a later period than the intestinal muscle, differentiates into a sphincter ductus choledochus, a sphincter ductus pancreaticus, and a sphincter ampullae, if an ampulla persists. At the point where the ducts enter the duodenum, this sheath receives auxiliary fibers from the intestinal muscle which tie the ducts to the intestine and assist in the erection of the papilla (**Fig. 4**). Because of this connection, some authors have denied the independence of the sphincter. It can act independently, however, and its spasms may result in one type of biliary colic. Physiologically its most significant component is the sphincter ductus choledochus. Tonic contraction of this sphincter causes filling of the gallbladder.

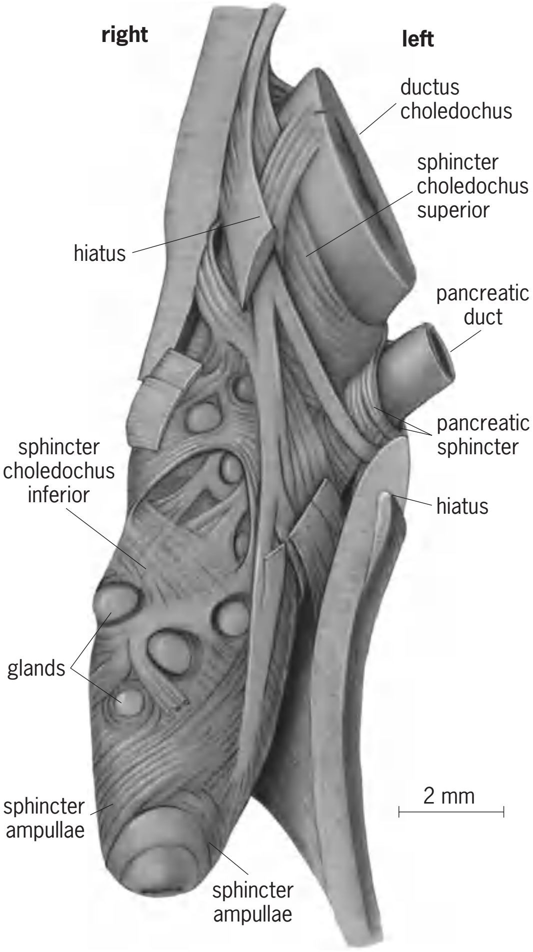


Fig. 4. Human sphincter of Oddi; upper and lower margins of window in muscle layer of duodenum through which bile and pancreatic ducts enter major duodenal papilla (Fig. 1). (*After E. A. Boyden, Surg. Gynecol. Obstet., 104(6):641–652, 1957*)

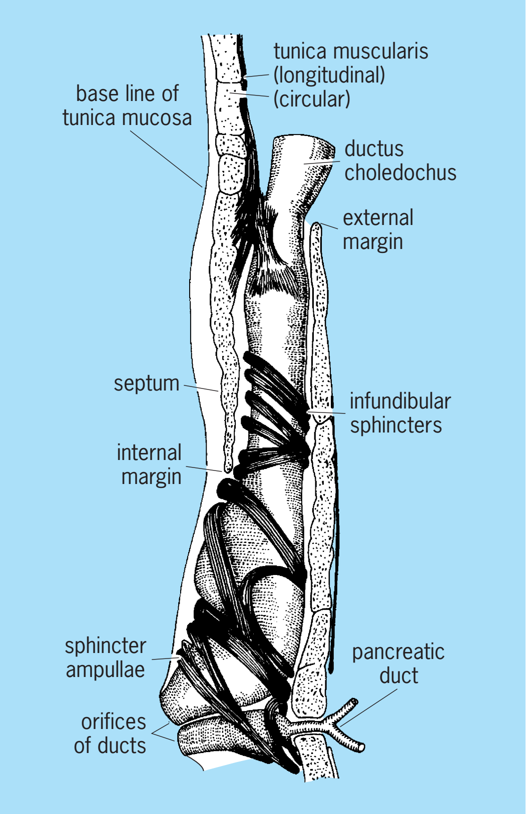


Fig. 5. Choledochoduodenal junction through wall of duodenum in newborn puppy. (*After E. P. Eichhorn and E. A. Boyden, Amer. J. Anat., 97:431–459, 1955*)

In other mammals the anatomy of the junction has been worked out in only a few species, namely the cat, dog, chimpanzee, and monkey. The first two differ from primates in that the bile duct passes through a funnel in the intestinal muscle before it enters the subterminal area where the special sphincters lie (**Fig. 5**). Therefore, bile flow in such species is regulated partly by control of intestinal activity. Experimental findings cannot be transferred from such species to humans without reservations.

Microscopic structure. The wall of the gallbladder consists of five layers: (1) a mucosal layer which lines the hollow viscus; (2) a vascular lamina propria (also named submucosal layer); (3) a fibromuscular layer containing entwined loose connective tissue and smooth muscle fiber bundles; and, where the organ is not attached to the liver, (4) a subserosal connective layer, rich in nerves, blood, and lymph vessels, (5) a serosal layer comprising a mesothelium or peritoneum. In areas where the gallbladder wall contacts the liver tissues, the subserosal layer is referred to as the adventitial layer. The mucosal layer is a simple columnar epithelium consisting of tall epithelial cells (pale and dark cells) and the subjacent lamina propria which contains a rich vascular supply, including fenestrated capillaries and lymphatics.

Both epithelium and lamina propria constitute a mucous membrane. In addition, like the intestinal lining, the mucosal epithelial cells contain lysosomal bodies and secrete mucus glycoproteins which cover the microvilli and contribute to protecting the cell surface from the potential detergent actions of the bile. This membrane concentrates the bile 7–10 times. If the membrane is diseased, these functions are lost or impaired; and the condition may be diagnosed by the Graham-Cole test, which depends on the concentration by the gallbladder of radiopaque substances administered orally. The surface area of the gallbladder mucous membrane is increased grossly by rugose folds and microscopically by countless minute cellular processes, the microvilli, each coated by a rich glycocalyx (**Fig. 6**). The surface lining epithelium is responsible for removing water and inorganic salts from the bile. These substances are transported into the basolateral spaces, then passed into the vascularized lamina propria which fills the rugae, and are evacuated by the subserosal or adventitial drainage. A rich nerve plexus originating from intramural ganglia provides numerous intraepithelial nerve endings. The surface epithelium invaginates deeply into the lamina propria and contributes to formation of pocketlike invaginations and extended rugae or folds. However, in case of inflammation, the latter may extend into and through the fibromuscular layer as pseudodiverticula of Rokitansky-Ashoff. Small mucous glands are found only in the neck of the gallbladder and within the wall of the cystic duct.

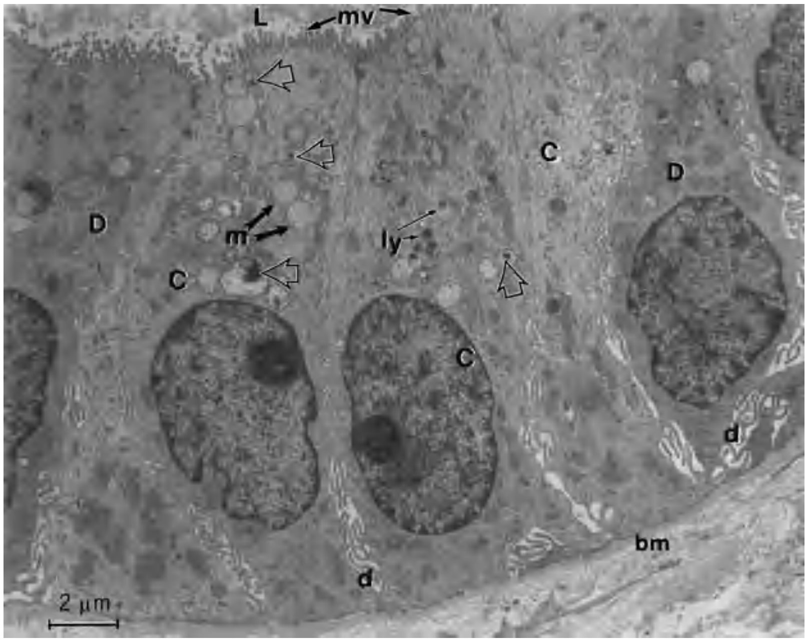


Fig. 6. General aspect of a normal epithelium from a sample of human cholecystic gallbladder. Columnar undifferentiated dark (D) and differentiated clear (C) cells produce mucus secretory granules (m) in their Golgi and subapical zones. Most granules contain proteinaceous electron-dense cores (large open arrows). Note the apical surface, lining the organ lumen (L), covered by numerous microvilli (mv); d, basolateral digitations; bm, basement membrane; ly, lysosomal bodies. (*After J. Gilloteaux et al., Amer. J. Anat., 184:323–335, 1989*)

Comparative physiology. In humans, evacuation of the gallbladder is accomplished by a trigger mechanism which is set off by the presence of fatty foods, meat, and hydragogue cathartics in the duodenum and upper jejunum. Absorption of these substances by the mucous membrane results in the release of cholecystokinin (CCK), a hormone which rapidly circulates in the bloodstream and simultaneously produces contraction of the gallbladder and relaxation of the sphincter of Oddi. The most effective food is egg yolk which contains certain *l*-amino acids. By calculating the changing volumes of the gallbladder in a series of x-rays, it has been shown that the first phase of contraction is the most important, because two-thirds to three-fourths of the contents is expelled within the first 40 min (**Fig. 7**). Resorption of bile salts by the intestine stimulates secretion of bile for hours after a meal. Before puberty, the male gallbladder empties faster than the female but it is slower afterward. Pregnancy and pernicious anemia in females markedly slow the rate. In peptic ulcer patients the rate of emptying is accelerated because initial food escapes faster from the stomach. After section of both vagus nerves, the gallbladder merely resumes its normal rate. For this and other reasons, it is believed that evacuation of the biliary tract in humans is controlled primarily by hormones.

In laboratory mammals the situation is more complicated. In animals below primates, gallbladder and sphincter are under control of both hormones and the autonomic nervous system. In cats contraction of the gallbladder can be inhibited within 1 min by inducing a spasm of the gut. The resulting reflex, from gut to gallbladder, can be blocked by sectioning the splanchnic nerves. Comparable experiments with human beings, in which medical students swallowed electrodes into the duodenum, failed to inhibit the gallbladder. Section of the right vagus in cats triples the time required to empty the gallbladder. Partial vagotomy in humans does not reduce the rate of emptying, and total vagotomy merely returns it to normal.

Several human clinical and comparative studies confirm and complete the aforementioned observations because the gallbladder appears to possess a rich ganglionated nerve plexus where circulating hormones, inflammation-related factors, and several co-localized neuropeptides can affect or modulate the actions of the acetylcholine release of the main vagal preganglionic gallbladder innervation. Humoral cholecystokinin has been confirmed to be the main postprandial stimulus for gallbladder contraction; it can readily access these gallbladder ganglia since there is no blood-ganglionic barrier. As a result of this lack of barrier, cholecystokinin can act presynaptically within the gallbladder ganglia, facilitating the acetylcholine release which in turn initiates gallbladder contraction via stimulation of the intramural intrinsic nerve plexus on the organ musculature. Norepinephrine and prostaglandin E_2 are known to antagonize the acetylcholine action.

The sphincter of Oddi also displays a ganglionated nerve plexus which contains 5-hydroxytryptamine (5-HT) neurons and fibers, suggesting that 5-HT plays a role in the sphincter regulation. The recent findings that duodenal sensory neurons project to the

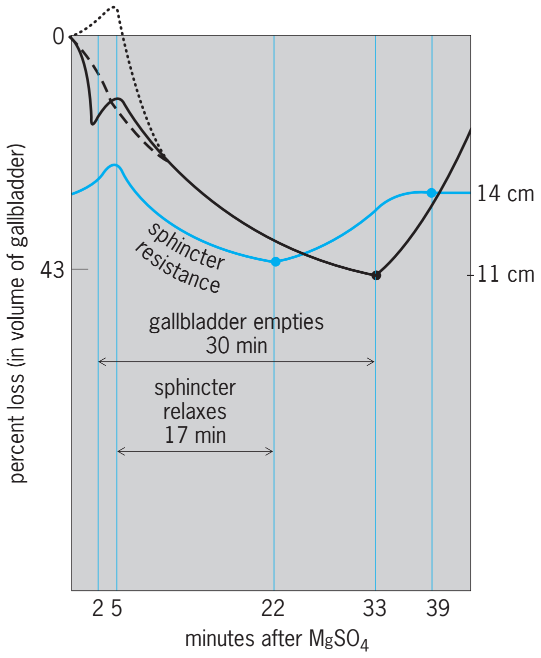


Fig. 7. Reaction of gallbladder and sphincter of Oddi to magnesium sulfate in humans. A mean drop from 14 to 11 cm Hg in pressure within common duct in 12 cholecystectomized patients; 43%, mean percent of bile expelled by gallbladder in 10 male students. When egg yolk was given, the times remained the same, but sphincter resistance fell from 14 to 7 cm of water, and gallbladder expelled 71% of its contents. (*After E. A. Boyden, G. S. Bergh, and J. A. Layne, Surgery, 13:723–733, 1943*)

sphincter of Oddi ganglia made plausible that these connections may detect the postprandial release of cholecystokinin from the mucosa and therefore provide the signal to decrease sphincter of Oddi resistance, and hence facilitate the flow of bile into the duodenum.

Functions in different species. Animals without a gallbladder, such as the pocket gopher, pigeon, horse, and rat, do not compensate for its absence by secreting a more concentrated bile. However, the liver of such animals secretes a relatively large quantity of bile. Conversely, animals secreting a relatively small quantity of bile, such as humans, the mouse, dog, and cat, have gallbladders which concentrate well. The gallbladders of the sheep, goat, cow, and pig appear to manifest little, if any, power of concentrating the bile. Correspondingly, the sphincter offers little or no resistance to the flow of bile. In humans and dogs, the gallbladder has both a storage and pressure-regulatory function; in the rabbit, sheep, cow, goat, and pig, only a pressure-regulatory function; and in still others, such as the guinea pig, no apparent function because its removal results in no grossly demonstrable changes. *See* DIGESTIVE SYSTEM.

Edward A. Boyden; Jacques Gilloteaux

Bibliography. P. W. Brunt and A. E. Read (eds.), *The Liver and Biliary System*, 1984; W. Ganong, *Physiology: A Study Guide*, 2d ed., 1987; J. Gilloteaux, The biliary tract, the gallbladder and gallstones, *Microsc. Res. Tech.*, 38(6) and 39(1), 1997; A. P. Gokin et al., *Amer. J. Physiol.*, 271:G876–G883, 1996, and 272:G1365–G1371, 1997; K. Hillsley and G. M. Mawe, *Amer. J. Physiol.*, 275(5, pt 1):G1018–G1027, 1998; W. P. Hopman et al., *Gastroenterology*, 89:1242–1247, 1985; A. S. Romer and T. S. Parsons, *The Vertebrate Body*, 6th ed., 1986; K. Schmidt-Nielsen et al. (eds.), *Comparative Physiology*, 1978.

Gallbladder disorders

Although not essential to life or health, the gallbladder is the site and source of appreciable suffering and disease in humans. With its cystic duct, the gallbladder constitutes a blind-ended, lateral extension of the common bile duct. It is roughly the size and shape of a small pear, and is adherent to the outer surface of the liver. Besides acting as a reservoir for bile, the gallbladder concentrates and otherwise alters the composition of bile. *See* GALLBLADDER.

Gallstones. Gallstones are round, oval, or faceted concretions formed within the gallbladder from the salts and pigment of bile. The size may be upward of 1 in. (2.5 cm) in diameter. Such stones may also be formed in any of the bile ducts within or outside the liver, but the incidence in the bile ducts is low compared to the number originating in the gallbladder. The occurrence of gallstones increases with age, and 20–30% of all elderly adults have one or more such stones. They are particularly common in women, blacks, and individuals with diabetes. Pure stones account for 10% of all gallstones, and are composed of cholesterol, calcium bilirubinate, or calcium carbonate, in decreasing frequency. Mixed types make up 80% of all gallstones and are most often associated with cholecystitis. The remaining 10% are the combined forms that usually begin as pure stones, with other basic materials being subsequently deposited.

Although the mechanism and reason for their formation are not clearly understood, the major predisposing factors are stasis (prolonged retention of bile in the gallbladder), abnormal composition of the bile (excessive amounts of cholesterol, bilirubin, or calcium), and infection. Passage of a gallstone through the ducts into the duodenum usually produces severe pain, called biliary colic. If a stone causes obstruction of the ducts, the result may be damage to the liver, pancreas, biliary system, and related structures either directly or through concomitant inflammation. Gallstones are rare in animals, although they have been found in nearly all species, especially in cattle. *See* BILIRUBIN; CHOLESTEROL; CIRRHOSIS.

Cholecystitis. Cholecystitis, or inflammation of the gallbladder, is a common disease in humans; as many as 70% of all elderly patients show anatomical evidence of having had one or more attacks. It is nearly always associated with gallstones and is particularly common in obese middle-aged women. It is rare in animals. Most cases are thought to be the result of chemical irritation caused by excessively concentrated bile, which is in turn the result of partial or complete obstruction to the outflow of bile. The gallbladder that is irritated and inflamed in this fashion is susceptible to secondary bacterial infection, and this in turn may spread throughout the bile ducts into the liver tissue. The inflamed gallbladder is usually enlarged because of edema of its wall and the presence of pus in its lumen. Perforation may occur, leading to the development of peritonitis or to abnormal communications with the intestine. Prolonged or recurrent episodes of inflammation result in chronic cholecystitis, characterized by thickening and scarring of the wall, contraction, and impairment of normal function. The latter may be detected clinically by administering a radiopaque dye which in a normal gallbladder is concentrated sufficiently to produce a shadow on the x-ray film. The absence of such a shadow indicates a failure of this concentrating ability and suggests past or present disease.

Tumors. Malignant tumors are less common in the gallbladder than in several other parts of the intestinal tract, such as the stomach, colon, and pancreas, but make up as much as 3% of all malignant tumors in humans. They are almost invariably associated with the presence of gallstones. Because they produce little in the way of symptoms, and because they very soon invade the liver, these tumors are rarely curable by surgical therapy at the time they are discovered. Similar carcinomas arise in the common bile duct and in the ampulla of Vater, but there the prognosis is somewhat better. In the large bile ducts, tumors usually cause biliary obstruction and jaundice early in their course and may be surgically resectable if the formidable technical problem of reestablishing biliary and pancreatic drainage can be surmounted. Benign tumors of the gallbladder and ducts are rare in humans, and in animals both benign and malignant tumors of these sites are extremely uncommon. *See* LIVER DISORDERS.

Milton R. Hales

Bibliography. T. C. Jones and R. D. Hunt, *Veterinary Pathology*, 6th ed., 1997; L. Okoliesanyi et al. (eds.), *Assessment and Management of Hepatobiliary Disease*, 1987; M. H. Sleisenger and J. S. Fordtran, *Gastrointestinal Disease: Pathophysiology, Diagnosis, and Management*, 4th ed., 1989; S. Sherlock, *Diseases of the Liver and Biliary System*, 10th ed., 1997.

Galliformes

A large order of birds containing the gallinaceous, or chickenlike, birds. They are found worldwide, although several of the subgroups have a more limited distribution. Part of the superorder Neognathae, the galliforms are closely related to the anseriforms and are often placed with them in a separate group, the Galloanserae. The gallinaceous birds include the most important domesticated birds, the chicken (*Gallus gallus*) and the turkey (*Meleagris gallopavo*); the rapidly breeding Japanese quail (*Coturnix japonica*) is an important species for much scientific work. The hoatzin of South America is often classified as a galliform suborder, but here it will be treated as a distinct order, the Opisthocomiformes. *See* ANSERIFORMES; AVES; OPISTHOCOMIFORMES.

Classification. The Galliformes are divided into two suborders with two families.

Order Galliformes
  Suborder Cracoidea
    Family: Megapodiidae (megapodes or brush turkeys; 7 species; Australasia to southeast Asia)
      Cracidae (curassows, guans, chachalacas, 49 species; New World, mainly Neotropics)
  Suborder Phasianoidea
    Family Phasianidae (grouse, pheasants, quail, guinea fowl, turkeys; 212 species)
      Subfamily: Tetraoninae, (grouse; Holarctic; temperate regions of the Northern Hemisphere)
        Phasianinae (pheasants, peafowl, chicken; Old World to Australia)
        Odontophorinae (New World quail)
        Numidinae (guinea fowl; Africa)
        Meleagidrinae (turkeys; Mexico to North America)

Although the galliforms are usually divided into two suborders, Cracoidea and Phasianoidea, and the Megapodiidae and Cracidae are considered primitive for the entire order, these conclusions must still be documented carefully.

Characteristics. The fowls are small to large birds with a short, arched bill for pecking seeds, insects, and vegetable matter. Their wings are short and rounded, and their flight muscles are large but myoglobin-poor (white); hence they fatigue quickly. Their flight is explosive but of short duration and is used mainly for purposes of escape. Although the fowls fly infrequently, no Recent species is flightless. Their legs and feet are strong, as they spend most of the time on the ground in search of food. Males of many species, especially pheasants, have colorful plumage and elaborate wing and, especially, tail feathers that are used in courtship displays. The train of the male peafowl, with its many complex ocelli (round, light-sensitive eyespots) and lacy appearance, is only an extreme of the trend seen in other species of galliform birds; moreover, this train is composed of upper covert feathers, not tail feathers. Courtship and pair bonding are brief; several species court communally in courtship display areas known as leks. The females are responsible for incubation and care of the young, which are downy and leave the nest immediately after hatching.

Cracoidea. The cracids (chachalacas and curassows) are restricted to South and Middle America and are the most arboreal members of the order; they have never hybridized with members of other families and may represent the most primitive group of galliform birds.

The megapodes (mound birds) do not build a nest or incubate the eggs. Rather, the eggs are buried in a mound of decaying vegetation, in warm sand, or in

White-tailed ptarmigan (*Lagopus leucurus*), Glacier National Park, Montana. (*Photo by Gerald and Buff Corsi, © 2001 California Academy of Sciences*)

earth heated by thermal vents. The young develop extensively, hatch fully feathered and self-sufficient, and can fly soon after hatching. The megapodes of Australia, and New Guinea to the Philippines and the Nicobar Islands are the only birds that do not exhibit parental care.

Phasianoidea. Grouse are Holarctic with colorful fleshy areas and pouches on the head and neck that are used in courtship. Some species eat conifer needles during the winter and/or burrow in the snow during the night to conserve heat. The ptarmigans (*Lagopus*), which have a brown summer plumage, have evolved a special white winter plumage that serves as camouflage, a characteristic that is unique in birds (*see* **illustration**).

Found in forests, turkeys are large galliforms with dark plumage that is metallic-colored in males. The skin of the head and neck is bare, and in males becomes red and blue during courtship. Males have a breast "brush" of unique specialized epidermal appendages, not of modified feathers.

The common quail of Eurasia (*Coturnix coturnix*) is notable because it is one of the very few migratory members of this order and it is differentially toxic to humans, which has been known since biblical times but is still poorly understood.

Fossil record. The earliest galliform fossil is *Gallinuloides* from the lower Eocene of Wyoming and may be an early cracid. A scattered fossil record of younger forms exists for all major groups of galliforms, but it does not clarify the relationships or evolutionary history of the order. Megapode fossils are interesting in that several giant, flightless forms are known from the Quaternary, including *Sylviornis* from New Caledonia, which was originally believed to be a ratite bird. *See* STRUTHIONIFORMES.

Economic significance. The galliforms have been one of the most economically important groups of birds to humans. Almost all species have been hunted for food, and many are still important as game birds for hunters. Most species are kept in captivity by aviculturists and are perhaps the most important group of domesticated birds, particularly chickens and turkeys. Chickens were domesticated from red jungle fowl by primitive humans in southeast Asia, rapidly spread over the Old World, and arrived in the New World with European explorers. Divergent breeds of chickens have been developed for egg and meat production. The wild turkey, which is found throughout most of the United States and Mexico and was domesticated by Mexican Indians in pre-Columbian times, was brought to Europe by early explorers. It is used exclusively for meat production. The Indian peafowl (*Pavo cristatus*) has been kept in captivity as an ornamental bird, and a white variety has been developed. One species of guinea fowl, *Numida meleagris*, has been domesticated on a much smaller scale. The ring-necked pheasant (*Phasianus colchicus*) of Asia has been introduced in many parts of the world as a game bird. Several other species have also been introduced as game birds, but with less success. Some domesticated gallinaceous birds have escaped and established themselves as feral birds.

Many species of galliforms have become greatly reduced in numbers and some threatened because of overhunting and habitat destruction. Some species, especially the forest pheasants of Asia, are seriously endangered. *See* ENDANGERED SPECIES.

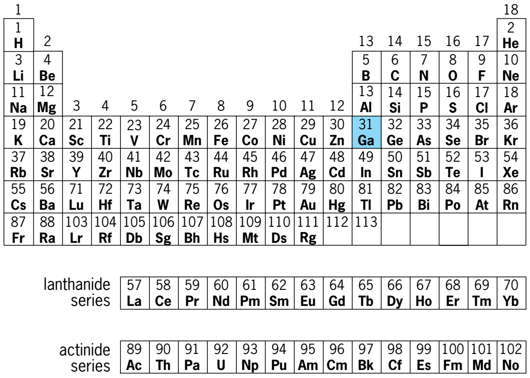
Walter J. Bock

Bibliography. T. M. Crowe, The evolution of the Guinea-fowl, *Ann. S. Afr. Mus.*, 76:43–136, 1978; J. Delacour and D. Amadon, *Curassows and Related Birds*, 2d ed., Lynx Edicions, 2004; P. A. Johnsgard, *The Grouse of the World*, University of Nebraska Press, 1983; P. A. Johnsgard, *The Pheasants of the World: Biology and Natural History*, 2d ed., Smithsonian Institution, 1999; P. A. Johnsgard, *The Quail, Partridges, and Francolins of the World*, Oxford University Press, 1988; D. N. Jones, R. W. R. J. Dekker, and C. S. Roselaar, *The Megapodes*, Oxford University Press, 1995; S. Madge, S. McGowan, and P. McGowan, *Pheasants, Partridges and Grouse*, Princeton University Press, 2002.

Gallium

A chemical element, Ga, atomic number 31, atomic weight 69.72. Gallium is a member of group 13 and the fourth period of the periodic table (IUPAC). *See* PERIODIC TABLE.

The major commercial sources of gallium are bauxite, containing gallite ($CuGaS_2$), and zinc and germanium sulfides. Normal ore-grade deposits usually



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

contain substantially less than 0.1% gallium. In the United States the bauxite deposits in Arkansas and the zinc deposits in Oklahoma are the main sources of domestic production. Much of the gallium used in the United States is imported from Switzerland and Germany, with lesser amounts from Canada and France. *See* BAUXITE.

Gallium is a unique element in that it possesses the largest liquid range of any element. Its normal freezing point of 29.78°C (85.60°F) is lower than any metal except mercury and cesium. Its boiling point is in the vicinity of 2420°C (4388°F), although there is some uncertainty owing to the reactivity of gallium with the container material at this temperature.

The valence-electron notation of gallium corresponding to its ground-state term is [Ar, $3d^{10}4s^24p^1$], which accounts for the maximum oxidation state of III in its chemistry. Compounds of formal oxidation state II and I are also known.

Approximately 95% of the gallium consumed in the United States and presumably in the world is used in the electronics industry. Minor quantities have been used or studied for use in thermometers, low-melting solders, as a heat-transfer fluid, in arc lamps, batteries, vanadium-gallium superconductors, and in catalyst mixtures.

The most important gallium semiconductors are gallium arsenide (GaAs) and gallium phosphide (GaP). The magnitude of the energy gap in GaAs favors its use in transistors. The electron mobility in GaAs is very much higher than the hole mobility; in contrast, the electron and hole mobility in GaP are of similar magnitude and very much lower than in GaAs. By doping with the appropriate elements, these properties can be altered. Electron transport (*n*-type) GaP semiconductors are used in rectifiers, hole transport (*p*-type) in light sources and photocells. *n*-Type GaAs semiconductors are used in injection lasers and *p*-type GaAs in electroluminescent transistors. *See* HOLE STATES IN SOLIDS; SEMICONDUCTOR.

GaN is prepared by the reaction of metallic gallium or Ga_2O_3 at elevated temperature with ammonia, and the other semiconductors by direct reaction with the elements or Ga_2O at high temperature. *See* ALUMINUM; INDIUM; THALLIUM. Edwin M. Larsen

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; A. J. Downs (ed.), *Chemistry of Aluminium, Gallium, Indium and Thallium*, 1993; W. F. Gale and T. C. Totemeier, *Smithells Metals Reference Book*, 8th ed., 2004; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; K. T. Tsen (ed.), *Ultrafast Physical Processes in Semiconductors: Semiconductors and Semimetals*, vol. 67, 2000.

Galvanizing

The generic term for any of several techniques for applying thin coatings of zinc to iron or steel stock or finished products to protect the ferrous base metal from corrosion; more specifically, the hot dipping that is widely practiced with mild steel sheet for garbage cans and corrugated sheets for roofing, sheathing, culverts, and iron pipe; and with fencing wire. During dipping, molten zinc reacts with the steel to form a brittle zinc-iron alloy. Control of temperature and the addition of aluminum reduce formation of the alloy, resulting in a more ductile coating. For marine use, magnesium is added.

An electrolytic process (also called cold galvanizing or electrogalvanizing) is also used for wire, as well as for applications requiring deep drawing. An alloy layer does not form, hence the smooth electroplated coating does not flake in the drawing die. *See* METAL COATINGS. Frank H. Rockett

Bibliography. R. M. Burns and W. Bradley, *Protective Coatings for Metals*, 3d ed., 1975; J. M. Hornsby, *Hot Dip Galvanizing: A Guide to Process Selection and Galvanizing Practice*, 1994; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 25, 1997; F. Lowenheim, *Modern Electroplating*, 3d ed., 1974.

Galvanomagnetic effects

Electrical and thermal phenomena occurring when a current-carrying conductor or semiconductor is placed in a magnetic field. The galvanomagnetic effects are closely related to the thermomagnetic effects. Both sets of effects yield important information on the band structure of metals and semiconductors and on the nature of the conductivity process. *See* THERMOMAGNETIC EFFECTS.

Let the electric current density j be transverse to the magnetic field H_z, for example, along x. Then the following transverse-transverse effects are observed: (1) Hall effect, an electric field along y, as shown in Eq. (1), where R is the Hall coefficient. According

$$E_y = Rj_xH_z \tag{1}$$

to experimental conditions, R may be the adiabatic or the isothermal Hall coefficient. (2) Ettingshausen effect, a temperature gradient along y, as in Eq. (2),

$$\frac{\partial T}{\partial y} = Pj_xH_z \tag{2}$$

where P is the Ettingshausen coefficient. Also the following transverse-longitudinal effects are observed: (3) Transverse magnetoresistance, an electrical potential change along x. (4) Nernst effect, a temperature gradient along x. *See* HALL EFFECT; MAGNETORESISTANCE.

Let the electric current density j be along H. Then, the most important effect is longitudinal magneto resistance, or an electrical potential change along H.

These effects arise from the Lorentz force on moving charges in a magnetic field. They are complicated by interaction between these charges and the lattice in which they move. For this reason, the signs of R and P may be negative or positive, depending upon composition and in many cases on temperature; the magnitudes show tremendous range.

Elihu Abrahams; Frederic Keffer

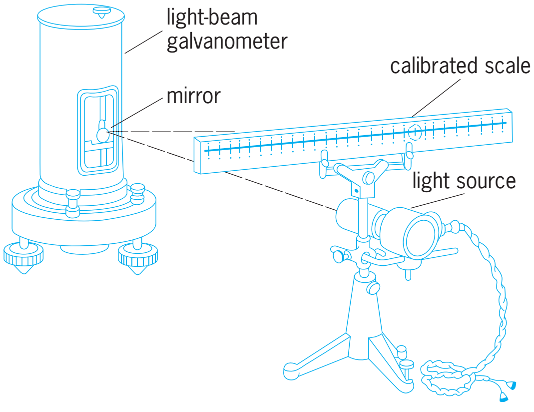
Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; R. E. Prange and S. M. Girvin (eds.),*The Quantum Hall Effect*, 2d ed., 1990; H. Weiss, *Structure and Application of Galvanomagnetic Devices*, 1969.

Galvanometer

A device for indicating very small, usually direct, electric currents. Although the deflection of a galvanometer results from current in the moving coil, the voltage in a closed circuit producing this current is frequently the quantity of interest to the user. In this mode, galvanometers are used to detect a null or an unbalanced condition in a bridge or potentiometer circuit. Galvanometers have played a pivotal role in the evolution of electrical measurements but electronic instruments that employ amplifying circuits to achieve sensitivities approaching the nanovolt level are now more frequently used instead.

D'Arsonval galvanometer. This is the most common type and is widely used. Its indicating system consists of a light coil of wire suspended from a copper or gold ribbon a few thousandths of an inch wide and less than 0.001 in. (0.025 mm) thick. The coil, free to rotate in the radial field between the shaped pole pieces of a permanent magnet, is attached to a pointer moving over a scale. In the most sensitive instruments, the "pointer" is a beam of light that is reflected off a mirror attached to the coil onto a fixed scale (see **illustration**). Current is conducted to and from the coil by the suspension ribbons. The torque which deflects the indicating element is produced by the reaction of the coil current with the magnetic field in which it is suspended, and is opposed by the twisting of the suspension ribbon. *See* MAGNETISM.

Sensitivity is optimized by choosing a coil resistance that matches that of the circuit as measured between the points to which the galvanometer is connected. Oscillations of the coil on applying or changing the current are damped out by the circuit resistance or by winding the coil on a short-circuited metal former.



Light-beam galvanometer, showing the detached scale and the light source as used in laboratory work. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

The sensitivity of galvanometers ranges up to 0.04 in. (1 mm) of deflection, on a scale 40 in. (1 m) distant from the mirror, for a current of 0.00001 microampere. Such a galvanometer may have a coil resistance of 800 ohms and a critical damping resistance of 100,000 Ω. The voltage response of this instrument amounts to 0.04 in./microvolt (1 mm/μV) at critical damping. A galvanometer by the same maker, designed for voltage sensitivity, has a coil resistance of 20 Ω, a critical damping resistance of 30 Ω, and a response of 0.04 in. (1 mm) for 0.05 μV in a critically damped circuit. It will be seen from these examples that a large response to current is associated with large coil resistance and high critical damping resistance, whereas voltage response is associated with low coil resistance and low critical damping resistance.

Portable galvanometers. Portable and some wall-type instruments have their coils suspended between upper and lower taut ribbons so that the motion is restricted to rotation about the coil axis. Such taut-suspension systems, however, are stiffer than those having the lower suspension loosely coiled, and such galvanometers are usually less sensitive. Pivoted galvanometers also have their motion confined to rotation. In both of these types the moving system is mechanically balanced around the axis of rotation, and the effect of mechanical disturbances is much reduced.

Ballistic galvanometer. The conventional galvanometer may also be used ballistically to integrate a transient current or voltage pulse of short duration. The electric charge in a capacitor may be measured by allowing it to discharge through a galvanometer and observing the magnitude of the first swing of the moving system from its rest position. A galvanometer's response to current of short duration is called its coulomb sensitivity. If the current pulse is short, the deflection of the moving system is proportional to the time integral of the current, that is, to the electric charge in the pulse. Similarly, a voltage pulse rising from a change in a magnetic field linking a coil may be integrated by a galvanometer. The time integral of voltage induced in the coil is proportional to the ballistic deflection of the galvanometer which is a measure of the total change in coupling between field and coil. This response defines the emf-time (or flux/linkage) sensitivity of the galvanometer. *See* AMMETER; CURRENT MEASUREMENT; ELECTRICAL MEASUREMENTS; ELECTROMAGNETIC INDUCTION; ELECTROMETER; VOLTAGE MEASUREMENT; VOLTMETER.

Forest K. Harris; Ronald F. Dziuba; Bryan P. Kibble

Bibliography. D. A. Bell, *Electronic Instrumentation and Measurements*, 2d ed., 1997; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; F. K. Harris, *Electrical Measurements*, 1952, reprint 1975; A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*, 1990; P. Kantrowitz et al., *Electronic Measurements*, 1979; L. M. Thompson, *Electrical Measurements and Calibration*, 2d ed., 1994.

Game theory

The theory of games of strategy can briefly be characterized as the application of mathematical analysis to abstract models of conflict situations. The first such models analyzed by the theory were parlor games such as chess, poker, and bridge. Since then, models arising from the behavioral sciences such as economics, sociology, and political science have been analyzed. Game theory is used in or closely connected to other areas such as linear programming, statistical decisions, management science, operations research, and military planning. In certain areas, the language and concepts of the theory are sometimes used even though the corresponding mathematics is not. *See* LINEAR PROGRAMMING; OPERATIONS RESEARCH.

Games in extensive form. Parlor games are specified by a list of rules. To be playable, a game must have the following properties: (1) There must be a way of starting the game. (2) There must be a well-defined list of legal (that is, permissible) moves for any possible situation that can be reached in the game. (3) At each move, exactly one of the players must be assigned to make the choice, or else the choice is made by a chance device (such as rolling dice or a spinning pointer). (4) After a finite number of moves, the game is terminated; a winner is declared or payments are exchanged among the players, or both. Any other conflict situation that satisfies mathematical axioms similar to these rules can also be considered a game and the analysis of game theory can be applied to it.

The players of a game are called persons, and such a person may actually consist of one or more people (for instance, in bridge the pairs of partners, east-west and north-south, each make up a player in the game; it is considered a two-person game). Thus a person in a game may be considered synonymous with a team. Chance moves occur when hands are dealt from a shuffled pack, dice are rolled, or pointers are spun. One says that all chance moves are allotted to the chance player—a fiction that is useful in abstracting properties of games.

When specified by a list of rules, a game is said to be in extensive form. For mathematical purposes, it is convenient to have games in normalized form, and for that, the idea of a pure strategy is needed.

A pure strategy for a player (not the chance player) is a complete list of choices of legal moves that he or she will make for every possible situation that can occur during the game. This is a much more complete list of decisions than that commonly called a strategy. For instance, in tic-tac-toe a small part of a single pure strategy would be, "If he moves to the center, I will move to the upper right-hand corner; but if he moves to the upper right-hand corner, I will move to the center . . . ; if he moves first to the center, and I to the upper right-hand corner, and he to the lower right-hand corner, then I move to the upper left-hand corner. . . . " The number of pure strategies that are possible in a game can be astronomical even for childish games such as tic-tac-toe, as any reader who tries to enumerate them will discover. Because of the enormous number of pure strategies, the actual applications of game theory even to parlor games have been severely limited by computational difficulties. Simplified versions of the games have been developed for which computations have been completely carried out.

Games in normalized form. After players have chosen pure strategies in a game, they need not physically play the game. Instead they could hand their strategies to a neutral person, or umpire, who could then carry out their instructions and make the moves they would have made. Their intuitively obvious idea leads naturally to the normalized form of the game.

Assume for the moment there are no chance moves in the game, that is, that there are n real players but no chance player. Denote by $s_1, s_2, \ldots, s_n$ specific pure strategies for players $1, 2, \ldots, n$, respectively. Given these, the game must be played in exactly one way and a unique outcome will result. Let $P_i(s_1, s_2, \ldots, s_n)$ be the monetary outcome to player i for this play of the game. (In case only a winner is announced, but no money is exchanged, it can be arbitrarily required that the losers pay one unit each and the winners divide the losers' payments equally, and in this way make a monetary assignment.)

Before the effect of the chance player can be introduced, the important concept of mathematical expectation must be explained. Suppose that $O_1, O_2, \ldots, O_k$ are mutually exclusive monetary outcomes of a chance event, and suppose further that they happen with probabilities $p_1, p_2, \ldots, p_k$ where $p_i > 0$ and $p_1 + p_2 + \cdots + p_k = 1$. Then the mathematical expectation E of the chance event is defined to be the sum $E = p_1O_1 + p_2O_2 + \cdots + p_kO_k$. *See* PROBABILITY.

As an example, suppose that one tosses a coin and is paid \$3 if heads turn up and loses \$1 if tails turn up. If it is assumed that the coin is balanced, each event will happen with probability 1/2 and the expectation is then $3(1/2) - 1(1/2) = 1$. Note that the expectation is not equal to either of the payoffs that can actually be obtained on a single toss of the coin. It can be interpreted as the average amount obtained if the experiment is repeated many times.

If there are chance moves in the game, a set of pure strategies, one for each player, will not determine a unique outcome of the game but merely a set of possible outcomes. These outcomes will be mutually exclusive and have probabilities depending on the chance moves associated with their occurrence. Hence in this case one can let $P_i(s_1, s_2, \ldots, s_n)$ be the expected payoff to player i for each $i = 1, 2, \ldots, n$.

Now the normalized form of a game is defined as the list of all expected payoffs to each player for every possible combination of pure strategies. In the case of two-person games it is most convenient to list these in tables called matrices.

Matching pennies. Two players R and C compare pennies, with R trying to match C. Each player can choose either heads or tails. Each has two pure strategies, "choose heads," H, or "choose tails," T. The game can be summarized as in tabulation (1). Here

	C chooses	
R chooses	H	T
H	1	−1
T	−1	1

(1)

the entries represent the payments by C to R. Thus, if both R and C choose heads, R matches C and receives one penny from C; but if R chooses heads and C tails, then R receives − 1 from C; that is, R pays one penny to C. The other two entries are determined similarly. The solution of this game is given below.

The Mississippi gambler. A passenger on a Mississippi river boat was approached by a flashily dressed stranger (the gambler) who offered to play the following game: "You take a red ace and a black deuce and I'll take a red deuce and a black trey; we will simultaneously show one of the cards; if the colors match I'll pay you and if the colors don't match you pay me; moreover if you play the red ace we will exchange the difference of the numbers on the cards; but if you play the black deuce we will exchange the sum of the numbers. Since if you lose you will pay me either \$2 or \$4 and if I lose I will pay you either \$1 or \$5, the game is obviously fair."

The proposed game is most easily summarized as in tabulation (2). The payoffs listed are those

	Gambler chooses	
Passenger chooses	Red 2	Black 3
Red 1	1	−2
Black 2	−4	5

(2)

received by the passenger. It is shown below that this game is not fair, and that the gambler has a substantial advantage.

Simplified poker. The deck used in simplified poker has only two kinds of cards, in equal numbers, high (H) and low (L). For instance, a bridge deck could be used with red cards high and black cards low. There are two players R and C and each player antes an amount a and is dealt one card which is his hand. It is easy to see that there are only four ways to deal the hands in the game, namely, (H, H), (H, L), (L, H), and (L, L), where R gets the first card of the pair and C the second.

After the deal, R has two alternatives, namely, to see or to raise by adding an amount b to the pot. If R sees, the higher hand wins the pot or equal hands split the pot equally. If R raises, C can fold or can call by adding b to the pot. If C folds, player R wins the pot (without revealing his hand). If C calls, the higher hand wins the pot or equal hands split the pot equally. These are all the rules.

Pure strategies for the players tell them exactly what to do in every situation. A pure strategy for R is to raise if a high card is dealt and to see if a low card is dealt. This can be abbreviated simply to raise-see. It is easy to see that R has four pure strategies, raise-raise, raise-see, see-raise, and see-see. Similarly, C has four pure strategies, fold-fold, fold-call, call-fold, and call-call.

Given a choice of a pure strategy for each player there are exactly four ways the game can proceed, depending on which of the four deals is made. if the mathematical expectation of each is computed, tabulation (3) results.

High		Fold	Fold	Call	Call
	Low	Fold	Call	Fold	Call
See	See	0	0	0	0
See	Raise	$\frac{3a}{4}$	$\frac{2a}{4}$	$\frac{a-b}{4}$	$\frac{-b}{4}$
Raise	See	$\frac{a}{4}$	$\frac{a+b}{4}$	0	$\frac{b}{4}$
Raise	Raise	$\frac{4a}{4}$	$\frac{3a+b}{4}$	$\frac{a-b}{4}$	0

(3)

Since a and b are positive numbers, R is better off choosing the fourth row rather than the second, and the third row rather than the first. Similarly, C, who is minimizing, prefers the third column to either of the first two. Therefore the maxtrix of tabulation (3) can be reduced to that of tabulation (4).

Play			Conservative	Bluffing
	High		Call	Call
		Low	Fold	Call
Conservative	Raise	See	0	$\frac{b}{4}$
Bluffing	Raise	Raise	$\frac{a-b}{4}$	0

(4)

In tabulation (4) the raise-see strategy is labeled as conservative for R and the raise-raise strategy as bluffing because these correspond to the usual notions. Similarly, call-fold is conservative and call-call is bluffing for C.

This game will be solved for various values of a and b in a later section.

Classification of games. A game is called zero-sum if, for every possible n-tuple of pure strategies $s_1, \ldots,$ s_n, the sum of the payoffs to all players is zero, that is, Eq. (5) holds. If this sum is not zero for some n-tuple

$$P_1(s_1, \ldots, s_n) + P_2(s_2, \ldots, s_n) + \cdots + P_n(s_1, \ldots, s_n) = 0 \quad (5)$$

of pure strategies, the game is called nonzero-sum. Examples of zero-sum games are matching pennies and poker, and also a wide class of games, called matrix games, discussed below. If money payments are assigned in bridge according to the points obtained by players, then bridge becomes a nonzero-sum game.

A game in extensive form is said to have perfect information if every player knows and remembers each move of each of her or his opponents as they are made. Examples of games with perfect information are board games such as checkers, chess, or backgammon. A game is said to have perfect recall if each player knows and remembers everything she did (but not necessarily what her opponent did). Poker is a game with perfect recall because each player knows what is in her own hand. However, bridge does not have perfect recall, because the personality of a player (= team) is divided between two actual persons and, for instance, north does not see what is in south's hand (except when south is dummy) even though they are members of the same team.

Games can now be classified according to the number of players they have, whether or not they are zero-sum, and whether or not they have perfect information or perfect recall. The games that will be discussed in more detail are (1) games in extensive form; (2) two-person zero-sum games or matrix games; (3) one-person nonzero-sum games, frequently called statistical games; (4) two-person nonzero-sum games; and (5) games involving three or more persons.

Signaling and behavior strategies. Some aspects of the theory can be studied only when games are specified in extensive form.

By a theorem of E. Zermelo and J. von Neumann, any game having perfect information has a pure strategy solution. Thus all board games can be solved in pure strategies. For chess this means that either (1) one of the players has a pure strategy that guarantees winning regardless of what the opponent does, or (2) each player has a strategy that will force the game to a draw or a win regardless of what the other player does. N. Dalkey has extended the theorem and has characterized the games that have pure strategy solutions.

H. W. Kuhn has shown that every game with perfect recall can be solved by means of behavior strategies. A behavior strategy is a set of probability weights assigned by a player to the alternatives at each of the choice points in the game. (One part of a behavior strategy for poker, for instance, would be to bet or fold with certain probabilities at a given turn in the game.) A consequence of this theorem is that poker can be solved by means of behavior strategies.

G. L. Thompson has shown that every game in extensive form can be solved by a combination of signaling strategies and behavior strategies. He has constructed a simplified model of bridge for which signaling strategies (partial pure strategies) are needed for the solution.

Matrix games. A matrix is a rectangular array of numbers such as those of tabulations (1) to (4). The matrices in tabulations (1) and (2) are called two-by-two matrices because they have two rows and two columns. The matrix in tabulation (3) is four-by-four. In general, a matrix is called m-by-n, if m is the number of its rows and n the number of its columns. Any two-person zero-sum game can be written in normalized form as a matrix game. The matching pennies and simplified poker games discussed above are examples.

Conversely, any matrix with numerical entries can be considered a two-person game between player R who chooses a row of the matrix and player C who chooses a column. After making their choices, C pays R the entry in the common row and column they have chosen (in case the entry is negative, R must pay C the amount).

A matrix then provides an abstract mathematical model for the study of two-person zero-sum games. What is meant by the solution to such a game? Consider the matrix of tabulation (6).

$$\begin{array}{c|cc} & \multicolumn{2}{c}{C \text{ chooses}} \\ \hline R \text{ chooses} & 1 & 2 \\ & -3 & 5 \end{array} \tag{6}$$

Player R (the row player) will receive the positive amounts of the matrix and must pay out the negative ones. This is the maximizing player. Similarly, C (the column player) must pay out the positive amounts and will receive the negative ones, so this is the minimizing player. In tabulation (6) player R would like to get the 5 entry, but to do so would have to choose the second row. And if R did this, player C would choose the first column and R would lose $3. Hence, R plays the first row and C, who wants to minimize payments, will still choose the first column. The result is that R will receive $1 from C.

The example of tabulation (6) is that of a strictly determined game, that is, a game in which there is a matrix entry that is simultaneously the minimum entry in its row and the maximum entry in its column. Such an entry is called a saddle-point entry of the matrix. Any matrix game that has a saddle-point entry has a solution, namely, that the players should choose rows or columns that contain such saddle-point entries. The value v of such a game is the saddle-point entry itself. In a strictly determined game the row player can be assured of at least v by playing an optimal strategy. To establish this, note that if this player chooses a row whose minimum entry is maximum, then the least payoff he can get is v, the minimum entry of this row. By a similar argument, the column player can be assured of not losing more than v by playing an optimal strategy.

However, the matrix games of tabulations (1) and (2) do not have saddle-point entries. How can such games be solved? Perhaps the matching pennies example of tabulation (1) suggests an answer. It is quite common for players of this game to randomize their choice of heads or tails by flipping the coin in the air before choosing. In this way they do not know themselves exactly what choice they will make but leave it to a chance device. Hence, their opponent cannot discover in advance their choice and cannot take advantage of such knowledge. Extensions of this idea give a solution concept to any game, strictly determined or not.

By a mixed strategy for a player in a matrix game is meant a set of probability weights, that is, nonnegative numbers that add up to 1. There are exactly as many probability weights as the player has pure strategies. Thus, in an m-by-n matrix game, a mixed strategy for player R is a set of numbers $p_1, p_2, \ldots, p_m$ with $p_1 \geqq 0$ and $p_1 + p_2 + \cdots + p_m = 1$. Similarly, the set of numbers $q_1, q_2, \ldots, q_n$, with $q_j \geqq 0$ and $q_1 + q_2 + \cdots + q_n = 1$, is a mixed strategy for player C. In using a mixed strategy, the player operates a chance device that chooses among the alternatives with probabilities equal to the corresponding weights. For instance, the device of flipping a coin in a matching pennies game has already been mentioned. In other games, the random choice of a pure strategy may be made by rolling dice, spinning pointers, or choosing random numbers.

A mixed strategy has the desirable property that a player's opponent cannot find out exactly what the player's choice will be, because the player does not know until the chance device has been operated. But even more is true. The use of a mixed strategy by a player introduces a chance move into the game and eliminates all personal moves. Hence, its effect in the game can be accounted for by computing each player's expected value, given a choice of mixed strategy for each. Let p stand for player R's mixed strategy and q for player C's. Let $E(p, q)$ be the expected payoff to R for these choices; then $-E(p, q)$ is C's expected payoff. Thus, if only the payoff $E(p, q)$ is considered, it is evident that R wants to choose p to maximize this amount and C wants to choose q to minimize it.

To analyze the game further, consider two modified games, one in which player C makes a choice of mixed strategy after R has announced a mixed strategy and the other in which R makes a choice after C's announcement. If C chooses after R, C will select a q that minimizes the result; because R knows that C will do this, R selects p initially to maximize the eventual result, that is, in this case p and q are chosen so that expression (7) is obtained. In the other game,

$$\max_p \min_q E(p, q) \tag{7}$$

in which R chooses after C, p and q will be chosen so that expression (8) is obtained. It is not hard to

$$\min_q \max_p E(p, q) \tag{8}$$

show that relation (9) holds for any matrix game. It is

$$\max_p \min_q E(p, q) \leq \min_q \max_p E(p, q) \tag{9}$$

obvious that, if R chooses a strategy according to the max-min rule, R is playing very conservatively and is being protected from the worst that C can do, and similarly for C and the min-max strategy.

The celebrated min-max theorem of von Neumann, proved in 1928, states that even more than relation (9) is true; namely, that for any matrix game there exist mixed strategy vectors $\mathbf{p}^0$ and $\mathbf{q}^0$ so that equality in relation (9) holds. In other words, relation (10) holds. It is this theorem that makes the study

$$\begin{aligned} E(\mathbf{p}^0, \mathbf{q}^0) &= \max_p \min_q E(p, q) \\ &= \min_q \max_p E(p, q) = v \end{aligned} \tag{10}$$

of games a mathematical theory rather than just a descriptive, intuitive theory. The strategies $\mathbf{p}^0$ and $\mathbf{q}^0$ are called optimal mixed strategies for the players R and C, respectively, and $v = E(\mathbf{p}^0, \mathbf{q}^0)$ is called the value of the game. The min-max theorem states, in words, that every matrix game has a value and at least one optimal strategy for each player. Thus every matrix game has a solution.

Many proofs of this important theorem exist. As it stands, it is an existence theorem, that is, it assures the existence of a solution to every matrix game, but it does not tell how to find it. Methods have been worked out for finding on electronic computers solutions to very large games.

Solution of examples. Although it is not possible here to go into the details of finding solutions to games, it is easy to verify solutions when stated.

Matching pennies. The solution to this game is for each player to choose pure strategies with equal probabilities, that is, $\mathbf{p}_1{}^0 = \mathbf{p}_2{}^0 = 1/2$ and $\mathbf{q}_1{}^0 = \mathbf{q}_2{}^0 = 1/2$. In other words the coin-flipping strategy usually employed by players of this game is actually optimal! The value of the game is 0 and it is called a fair game.

Mississippi gambler. Here the optimal strategies are $\mathbf{p}_1{}^0 = 3/4$, $\mathbf{p}_2{}^0 = 1/4$, $\mathbf{q}_1{}^0 = 7/12$, $\mathbf{q}_2{}^0 = 5/12$, and the value of the game is $-1/4$. In other words the passenger can expect to lose an average of a quarter each time the game is played!

Simplified poker. Two special cases for selected values of a and b will be solved.

1. $a = 4$ and $b = 8$. The matrix of tabulation (4) becomes tabulation (11), which is a strictly determined game. Each player should play conservatively, and the value of the game is 0.

	Conservative	Bluffing
Conservative	0	2
Bluffing	–1	0

(11)

2. $a = 8$ and $b = 4$. The matrix of tabulation (4) becomes tabulation (12). Here the value of the game is 1/2 meaning that the game is biased in favor of player R. Optimal strategies are that each player should bluff

	Conservative	Bluffing
Conservative	0	1
Bluffing	–1	0

(12)

half the time and play conservatively half the time. In ordinary parlance, bluffing is regarded as aggressive play, but note that, to play optimally, it is necessary to bluff half the time.

Statistical games. One of the first applications of game theory to another field was its use in the theory of statistical decision functions by the statistician Abraham Wald. In such statistical games, the row player is interpreted to be the statistician and the column player, Nature. The difference between Nature and a real-life player is that there is no known bias on her part. She may sometimes choose to help the statistician and sometimes not. Wald himself chose to apply the theory of games to this problem directly and gave the solution for the statistician as the minmax strategy in the matrix game. This solution has been objected to by others as being too pessimistic, and several other solutions have been proposed.

For example, in the Doctor's Dilemma, the statistician is a doctor who is testing two drugs *A* and *B* for the treatment of a certain disease. Drug *A* is known to be somewhat effective for treatment and never has been associated with aftereffects. Drug *B* has been spectacularly effective in certain cases but there have been undesirable aftereffects that, however, cannot be definitely attributed to the drug alone. The doctor analyzed the situation to be as in tabulation (13).

	Nature's choice	
Doctor's choice	*A* better than *B*	*B* better than *A*
Uses *A* rather than *B*	10	5
Uses *A* rather than *A*	−5	20

(13)

She assigned these payoffs by associating a positive monetary outcome to early recovery from the disease and a negative outcome to undersirable effects. How shall the doctor make her choice?

To protect herself from the worst that can happen using pure strategies, she would choose the first row, because it is the row whose minimum entry is maximum. If she permits herself mixed strategies, she would use the optimal strategy that selects the first row with probability 5/6 and the second row with probability 1/6. (For instance, she could roll a die and choose drug *B* if and only if an ace turns up.)

For many people the use of chance devices in making important decisions such as these seems repugnant. Yet the theory of games shows that, if the goal is to maximize mathematical expectation, and the analysis of the situation is correct, then such a decision rule is optimal.

Nonzero-sum games. By far the most satisfactory part of the theory of games consists of the zero-sum two-person cases, that is, in matrix games. Applications of the theory to such areas as economics, sociology, and political science almost invariably lead to many-person nonzero-sum games. Although no universally accepted theory has been developed to cover these games, many interesting and useful attempts have been made to deal with them.

When more than two persons are involved in a conflict situation, the important feature of the game becomes the coalition structure of the game. A coalition is a group of players who band together and, in effect, act as a new player in the game. There are two extremes to be considered. One is the noncooperative game in which such coalitions are banned by some means. Equilibrium-point solutions, discussed below, provide reasonably satisfactory solutions to such games. The other extreme is that in which all the players join together in a coalition to maximize jointly their total payoff. A game in which coalitions are permitted is called a cooperative game.

In the noncooperative game, each player is solely interested in his own payoff. By an equilibrium point in such a game is meant a set of mixed strategies $s_1, \dots, s_n$ such that Eq. (14) holds for each $i =$

$$P_i(s_1, \dots, s_i, \dots, s_n) = P_i(s_1, \dots, s'_i, \dots, s_n) \qquad (14)$$

$1, \dots, n$ for all strategies S_i of player i. What this means is that no player can, by changing strategy and assuming that the other strategies stay fixed, improve the payoff. By a theorem of J. Nash, every game has at least one equilibrium-point solution (commonly there are several).

For example, in the gas war, Jones and Smith own gasoline stations on opposite sides of the street. No other stations are nearby. There are only two prices they can charge, high and low, and each day they must decide which price they will charge. They are not permitted to change prices during the course of the day. Suppose their daily gross receipts are as in tabulation (15). In each pair of numbers enclosed in parenthesis the first number is Jones' receipts and the second number is Smith's receipts.

	Smith	
Jones	High	Low
High	(10, 10)	(6, 16)
Low	(16, 6)	(7, 7)

(15)

To interpret these numbers, observe that if both charge the high price, each receives $10. If both charge the low price (gas war), their income is cut to $7 for the day. Finally, if one charges the high price and the other low, the low-priced one can not only draw business away from the immediate competitor but also from elsewhere; and the volume of business yields $16, whereas the opponent gets only $6.

The unique equilibrium point is (7, 7) as can easily be checked. In other words the gas war prevails. If either player changes his strategy, he will only cut his own income, so that this solution is stable in the sense that neither can, by himself, profitably change to another strategy.

Cooperative solutions. To most people the equilibrium-point solution to the gas-war example seems irrational, because at the (10, 10) point each player is better off than at the equilibrium point. But the (10, 10) point is not stable, because either player can, by switching to the low price, improve the take from 10 to 16 (assuming that the opponent does not change). It is also true that at the (6, 16) or (16, 6) points the total amount the players take in is greater than the total amount they take in at the (10, 10) point, and so either of these points is preferable. Yet, again, neither of these points is stable in the above sense.

The von Neumann-Morgenstern cooperative solution to the gas-war game is that the players should join together and cooperate to maximize their total intake; they should divide it in such a way that each player gets at least 7, the amount that can be received in the noncooperative solution. In the example, each player would get 7 plus some part of the surplus profit of 8, but the theory does not show further how division of the profit should be made.

Nash and H. Raiffa have given rationale that indicates that the surplus should be divided equally and the net payoff to each player should be $11. The division of the surplus could be carried out by means of a payment of one player to the other or else by having Jones charge high and Smith low on alternate days, and on the other days let the reverse be true. Then their average payment will be $11 each.

M. Shubik and G. L. Thompson have proposed another solution to the game. Suppose that Jones is richer than Smith and suppose that at the (7, 7) point both players actually lose money. Because the game is played over and over again every day, one could imagine Jones making a threat to Smith as follows: "If you do not joint with me in charging the high price, the day after you lower your price I shall lower mine and keep it down until you are bankrupt." Because Jones has more capital than Smith, the threat is effective and, if carried out, would result in the closing of Smith's business. If Smith tried to make a similar threat, he would simply ruin himself, and his threat is therefore suicidal. Here Jones is able to enforce the (10, 10) point and has the dominant position, in what otherwise appears to be a symmetric game, simply by being richer.

Recently, W. F. Lucas has shown that not every *n*-person game has a solution in the sense originally stated by von Neumann and Morgenstern. Shapley and Shubik have given economic interpretations to this result. They are also applying the theory of games to a series of economic problems.

Simple games. An important class of *n*-person games for applications to political behavior are the so-called simple games. Each coalition in such a game can be either winning, losing, or blocking. For instance, a winning coalition may be a set of voters who can elect their candidate, or a group of lawmakers who can pass their bill. The players not in a given winning coalition form a losing coalition. Finally, a coalition is blocking if neither it nor the players not in it can enforce their wishes.

It is a common observation that the power of an individual in such a voting situation is not always proportional to the number of votes he controls. This results in part from personality differences, and in part from the voting structure itself. The power index, proposed by L. S. Shapley, assigns to each player an a priori number that indicates the power that the voting structure gives to that player. It is illustrated in the following example.

Green, Jones, and Smith are on a committee. On any issue Green can cast 1 vote, Jones 2 votes, and Smith 3 votes. If it takes 4 votes to carry a measure, what are the winning, losing, and blocking coalitions and what is the relative voting power of each man? The winning coalitions are (Green, Smith), (Jones, Smith), and (Green, Jones, Smith). Losing coalitions are (Green) and (Jones). Blocking coalitions are (Green, Jones) and (Smith).

Suppose that the men enter the committee room singly. A man is pivotal when his entry into the room changes the set of people in the room from a losing or blocking coalition into a winning coalition. There are 6 orders in which they can enter the room. The pivotal man's name is listed in **boldface**. Since smith is pivotal 4 out 6 times he is assigned power 4/6 = 2/3. Similarly, Green and Jones are pivotal once each and each are assigned power 1/6. Note that the sum of the powers adds up to 1. Note also that Smith controls only one-half of the votes but has two-thirds of the power. Also Jones controls 1/3 of the votes but has only 1/6 of the power:

Green, Jones, **Smith**
Green, **Smith**, Jones
Jones, Green, **Smith**
Jones, **Smith**, Green
Smith, **Green**, Jones
Smith, **Jones**, Green

This model can be used to compare the actual power of a voting member with his index as given above. Thus personality factors can be separated from voting structure factors by means of the model. Generalizations of the model have also been made.

Continuous games. If, in the normalized form of the game, each player is permitted to have a continuous range of pure strategies and the payoff function is permitted to be a function of the two real variables that range over each player's strategies, the result is a continuous game. The model most frequently used is to let player R choose a number x between 0 and 1 and C choose a number y between 0 and 1, and let the payoff function $f(x, y)$ be defined on the unit square. Such games are frequently labeled "games on the square" for obvious reasons. There exist such games without solutions, but with some reasonable assumption as to properties of $f(x, y)$, it can be shown that the min-max theorem holds for these games. The same theorem has been generalized still further by permitting the strategy spaces of the players to be the more general spaces of analysis.

Game-playing machines. One of the first applications of large electronic computers to numerical problems was in solving large matrix games. Several methods have been devised for finding such solutions. One such method is the simplex method. It has been programmed to solve games with $m + n \geqq 4000$. So-called decomposition methods can extend these methods to certain problems having thousands or even millions of variables and constraints. The principal application of the method is to solve linear programming problems, which can be shown to be equivalent to matrix games.

Computers have also been programmed to play board games such as checkers and chess. Strictly speaking they do not use the theory of games at all at present, but instead use some of the game sense of the people who devised the codes. *See* ARTIFICIAL INTELLIGENCE; DECISION THEORY; INFORMATION THEORY; MATRIX THEORY; STOCHASTIC PROCESS.

Gerald L. Thompson

Bibliography. J. W. Friedman, *Game Theory with Applications to Economics*, 2d ed., 1990; J. McDonald, *Strategy in Poker, Business and War*, 1989, reissued 1996; P. Morris, *Introduction to Game Theory*, 1994; M. J. Osborne and A. Rubinstein, *A Course in Game Theory*, 1994; M. Shubik, *Mathematics of Conflict*, 1983; P. D. Straffin, *Game Theory and Strategy*, 1993; S. Vajda, *Mathematical Games and How to Play Them*, 1992.

Gametogenesis

The production of gametes, either eggs by the female or sperm by the male, through a process involving meiosis.

Animals. In animals, the cells which will ultimately differentiate into eggs and sperm arise from primordial germ cells set aside from the potential somatic cells very early in the formation of the embryo. In some cases, the area of the uncleaved egg which will be included as the cytoplasm of the primordial germ cells is recognizable by its content of organelles, called polar granules. After the embryo gastrulates, prospective germ cells, which are distinguishable from prospective somatic cells by their staining properties, often migrate to sites such as the extraembryonic membranes, where they remain until the gonads are constructed by somatic cells. The primordial germ cells then migrate to and populate the gonads. Evidence indicates that they make no contribution to somatic tissues, and that if they are lost during migration no gametes can be formed.

The final products of gametogenesis are the large, sedentary egg cells, and the smaller, motile sperm cells. Each type of gamete is haploid; that is, it contains half the chromosomal complement and thus half as much deoxyribonucleic acid (DNA) as the somatic cells, which are diploid. Reduction of the DNA content is accomplished by meiosis, which is characterized by one cycle of DNA replication followed by two cycles of cell division. The primordial germ cells are designated either oogonia or spermatogonia when they have arrived in the gonads. In mammals, oogonia divide extensively until birth, while spermatogonial divisions continue throughout the life of the male. At the end of the gonial divisions, cells in both sexes enter first meiotic prophase, during which the chromosomes assume a filamentous appearance and homologous chromosomes pair and may exchange segments. In oogenesis a "lampbrush'' morphology often characterizes the chromosomes of the meiotic prophase. *See* CHROMOSOME; MEIOSIS.

The production of sperm differs from that of oo- cytes in that each primary spermatocyte divides twice to produce four equivalent spermatozoa (**Fig. 1**) which differ only in the content of sex chromosomes (in XY sex determination, characteristic of mammals, two of the sperm contain an X chromosome and two contain a Y). The morphology of sperm is highly specialized, with distinctive organelles forming both the posterior motile apparatus and the anterior acrosome, which assists in penetration of the oocyte at fertilization. *See* SPERMATOGENESIS.

The cytoplasm of the primary oocyte increases greatly during the meiotic prophase and often contains large quantities of yolk accumulated from the blood. Meiotic divisions in the oocyte are often set in motion by sperm entry, and result in the production of one large egg and three polar bodies (**Fig. 2**). The polar bodies play no role, or a very

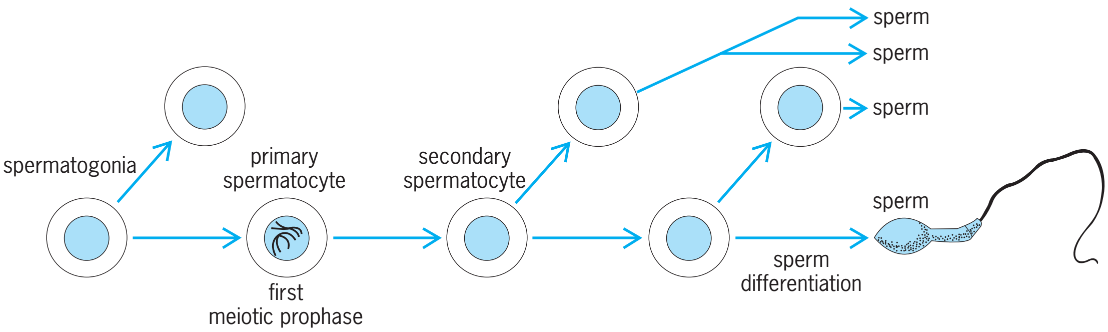


Fig. 1. Generalized diagram of spermatogenesis in animals.

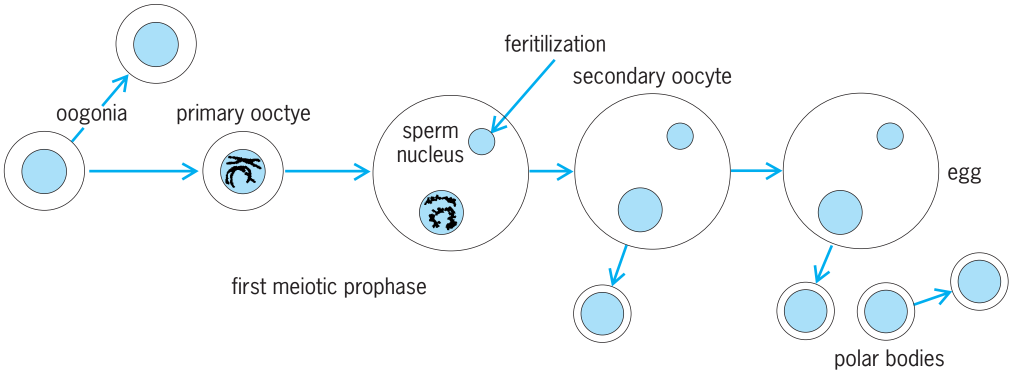


Fig. 2. Generalized diagram of oogenesis in animals.

subordinate one, in the formation of the embryo. *See* OOGENESIS.

After fertilization and the formation of the polar bodies, the haploid sperm and egg nuclei (pronuclei) fuse, thus restoring the normal diploid complement of chromosomes. *See* FERTILIZATION (ANIMAL); REPRODUCTION (ANIMAL). Spencer J. Berry

Flowering plants. Meiosis in flowering plants, or angiosperms, is essentially similar to that in animals. However, the cells produced after meiosis are spores, and these do not develop directly into gametes. Female spores (megaspores) and male spores (microspores) develop into gametophytes, that is, female and male haploid plants that bear within them the egg and sperm, respectively.

Female gametophyte. Development of the female gametophyte takes place within a developing ovule in the ovary of the pistil, the female structure of the flower (**Fig. 3**). One of the cells in the nucellus of the ovule differentiates into a megaspore mother cell, or megasporocyte, which undergoes the two meiotic divisions to form a linear tetrad of haploid (1*N*) megaspores. Three of the megaspores degenerate; the fourth enlarges and the nucleus divides

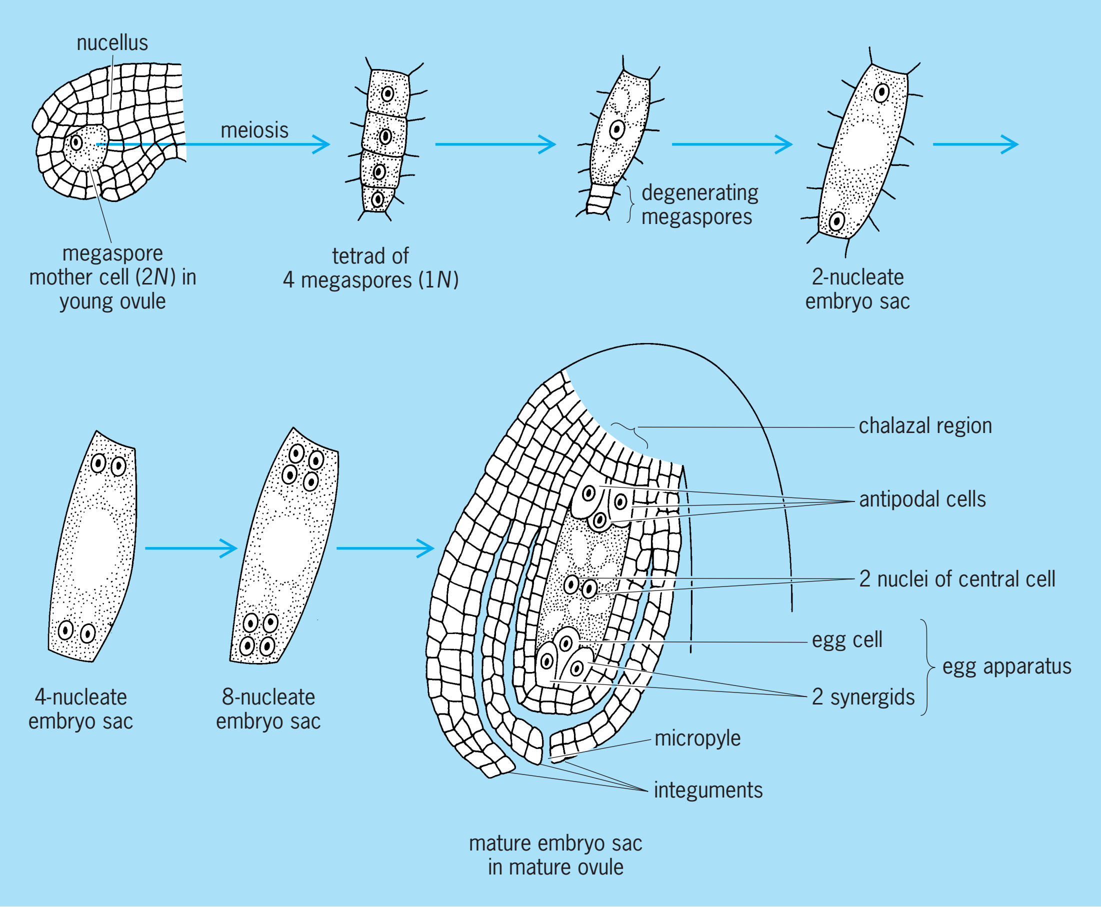


Fig. 3. Stages in the development of the embryo sac and female cells involved in fertilization in angiosperms.

mitotically, forming a two-nucleate embryo sac. Each nucleus divides mitotically, forming a four-nucleate embryo sac, and then each of these nuclei divides, giving rise to an eight-nucleate cell. At this stage, one nucleus from each group of four moves to the center of the cell to form a two-nucleate central cell. The cell is consequently organized into an embryo sac with seven cells through delimitation by cell membranes and cell walls around the remaining six nuclei with their associated cytoplasm. The embryo sac is now the female gametophyte. The three haploid cells at the chalazal end of the embryo sac are the antipodal cells, and these do not play any direct role in fertilization. The three haploid cells at the micropylar end compose the egg apparatus and consist of two synergids and an egg cell. The two nuclei of the large central cell may or may not fuse before fertilization, depending on the plant species. This is the most common type of embryo sac development found in flowering plants. Several other types of development and organization of the embryo sac, however, are known to occur in different plant species.

Male gametophyte. The development of the male gametophyte (**Fig. 4**) occurs within the anther, the male structure of the flower. Each microspore mother cell (microsporocyte) in the anther undergoes meiosis, producing a tetrad of four microspores which enlarge rapidly. The microspore nucleus divides mitotically but unequally to form the young male gametophyte, or pollen grain, which consists of two cells, the vegetative cell and the generative cell. The generative cell is small and lies within the cytoplasm of the vegetative cell. In the pollen of several plant species the generative cell divides in the pollen grain, producing two sperm cells. In most plants, however, the generative cell completes its division much later when the mature pollen grains are released from the anther and are deposited on the stigma of the pistil. After reaching the stigma, each pollen grain germinates, that is, extrudes a pollen tube, and the generative cell completes its division during the growth of the pollen tube but prior to the time it enters the embryo sac. The contents of the pollen tube are discharged into one of the synergids, and double fertilization occurs: One sperm cell fuses with the egg, giving rise to the zygote, while the other fuses with the central cell, giving rise to a 3*N* primary endosperm cell. Flowering plants are unique in nature in their requirement of double fertilization for the formation of a seed which consists of an embryo and the endosperm.

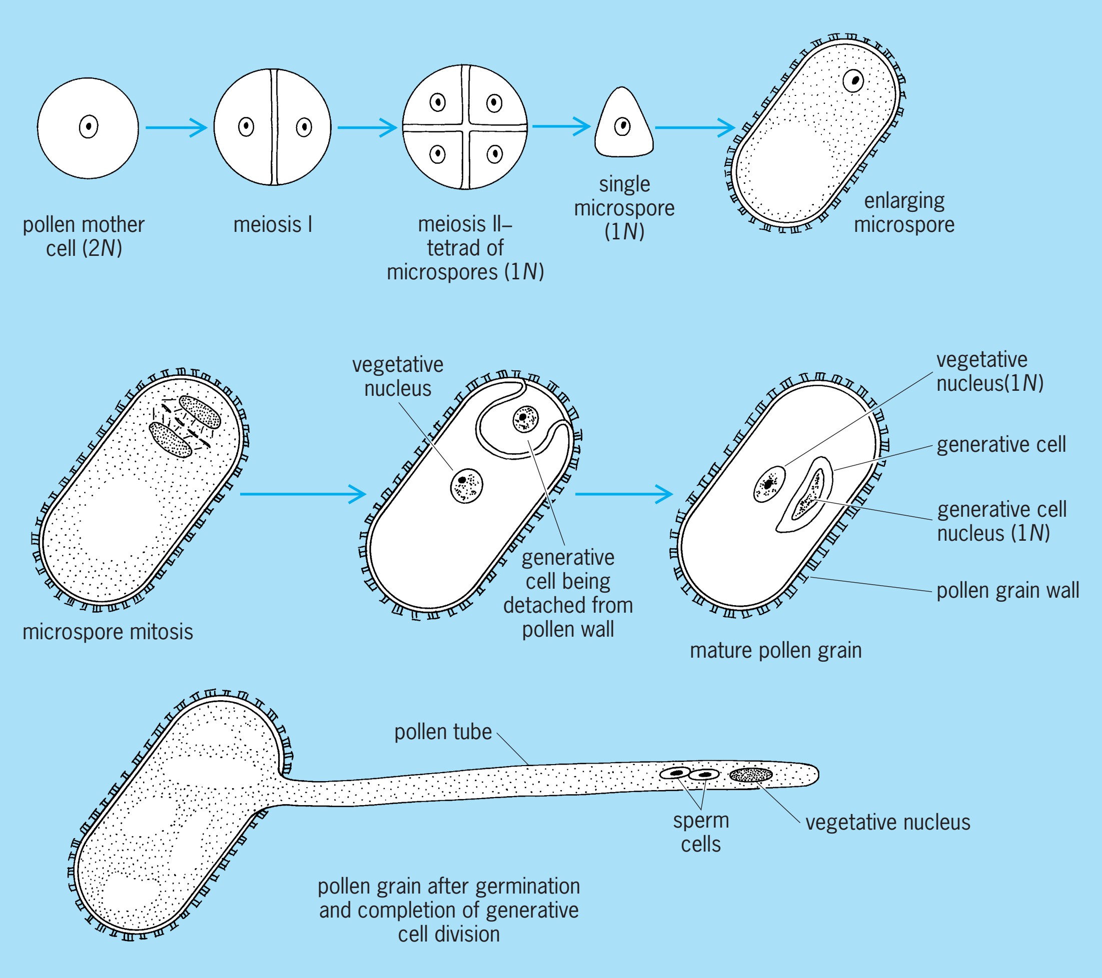


Fig. 4. Stages in the development of the pollen grain and sperm cells.

Other plants. Gametogenesis in angiosperms, the most advanced plants, has been considered in some detail, but there is a wide range in the details of development and structure of gametes among the different groups of plants.

Gymnosperms. In pines, for example, at the time of pollination, the male gametophyte consists of 4 cells, two prothallial cells, a generative cell and a tube cell. During pollen tube growth the generative cell divides to form a sterile cell and a spermatogenous cell. A short time before fertilization the spermatogenous cell divides to form two male gametes which are unequal in size. After meiosis during female gamete development, a single functional haploid megaspore is produced. The female gametophyte developing from the megaspore consists of about two thousand cells. Depending on the species of pine, within each female gametophyte 1 to 6 archegonia, or female gametangia, develop. Each multicellular archegonium produces only a single, nonmotile egg which at the time of fertilization is situated at the base of the archegonial canal. The pollen tube enters between the neck canal cells of the archegonium discharging its contents near the egg cell, enabling fertilization to occur. The larger of the two male gametes fuses with the egg cell. The smaller male gamete does not participate in fertilization.

Lower plants. In all lower vascular plants such as ferns, *Selaginella*, etc., no structure equivalent to the pollen tube is produced; the sperm are motile, flagellated cells that require free water to reach the passive, nonmotile egg. The mechanisms of gametogenesis and fertilization in plants are thus complex, exhibiting great variation between the different groups of plants. *See* DEVELOPMENTAL BIOLOGY; FERTILIZATION (ANIMAL); FLOWER; POLLINATION; REPRODUCTION (PLANT); SEED.

Joseph P. Mascarenhas

Bibliography. L. W. Browder, *Developmental Biology*, 3d ed., 1997; K. Esau, *Anatomy of Seed Plants*, 1977; A. S. Foster and E. M. Gifford, *Comparative Morphology of Vascular Plants*, 1974; P. D. Niewkoop and L. A. Sutasurya, *Primordial Germ Cells in the Chordates*, 1979; S. D. Russell and D. D. Cass, Ultrastructure of the sperms of *Plumbago zeylanica* L.: Cytology and association with the vegetative nucleus, *Protoplasma*, 107:85–107, 1981.

Gamma function

A particular mathematical function that can be used to express many definite integrals. There are, however, no significant applications where the gamma function by itself constitutes the essence of the solution. Instead it occurs usually in connection with other functions, such as Bessel functions and hypergeometric functions.

A special case of the gamma function is the factorial $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ (for example, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$). It is defined only for integral positive values of n. The factorial occurs, for instance, in the expansion

$$\exp z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

The binomial coefficient (N/n) can be expressed in terms of factorials as $N!/[n!(N-n)!]$. Many occurrences of the factorial are found in combinatorial theory (for instance, $n!$ is the number of permutations of n different elements), in probability theory, and in the applications of this theory to statistical mechanics. For large values of n, the factorial can be easily, although only approximately, computed with Stirling's formula,

$$n! \approx \left(\frac{n}{e}\right)^n (2\pi n)^{1/2}$$

This formula is in error by a factor, which lies, for all $n \geqq 1$, between 1 and $1/(11n)$. For $n = 10$, Stirling's formula gives approximately 3598695.6 ... instead of the accurate value $10! = 3628800$.

The gamma function can be considered as a certain interpolation of the factorial to nonintegral values of n. It is defined by the definite integral

$$\Gamma(z) = \int_0^\infty u^{z-1} e^{-u}\,du$$

This definition is applicable to all complex values of z with positive real part. The value of $\Gamma(n+1)$ coincides with $n!$ if $n = 1, 2, 3, \ldots$. The definition of $\Gamma(z)$ can be extended to all complex values of z by the principle of analytic continuation.

Other notations are in use: $z!$ or $\Pi(z)$ for $\Gamma(z+1)$. Sometimes the phrase factorial function is used for $z!$ or $\Pi(z)$ when z is arbitrary.

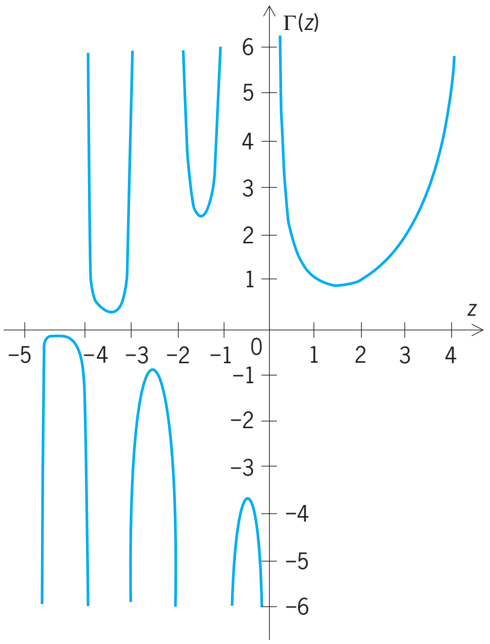
From its definition follow the fundamental properties of the gamma function:

1. $\Gamma(z+1) = z\Gamma(z)$ (difference equation or functional equation).
2. $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$, $\Gamma({}^1\!/_2 + z)\Gamma({}^1\!/_2 - z) = \pi/\cos(\pi z)$.
3. $\Gamma(z)\,\Gamma(z + {}^1\!/_2) = \pi^{1/2} 2^{1-2z}\Gamma(2z)$ (Legendre's duplication formula).
4. $\Gamma(z)$ is an analytic function in the complex z plane except for the points $z = -m$ ($m = 0, 1, 2, \ldots$), where it goes to infinity, as does $[(-1)^m m!\,(z+m)]^{-1}$, if z is close to $-m$.
5. As shown in the **illustration**, $\Gamma(z)$ is real for real values of z, $\Gamma(1) = 0! = 1$, $\Gamma(1/2) = \pi^{1/2}$.

In the angular domain $-\pi < \arg z < \pi$, there exists an expansion of the gamma function, namely,

$$\Gamma(z+1) = z^{z+1/2} e^{-z} (2\pi)^{1/2} \left(1 + \frac{1}{12z} + \frac{1}{288z^2} + \cdots\right)$$

which is not convergent but nevertheless useful for the numerical computation of $\Gamma(z)$ for large z. It gives approximate numerical values of the gamma function for large values of z if one takes only a finite number of terms into account. In particular for real positive values of z, the error is less than the last term taken into account. This expansion generalizes Stirling's formula.



Depiction of gamma function Γ(*z*) for real values of *z*.

Psi function. The logarithmic derivative of the gamma function $d \ln \Gamma(z/dz = \Gamma'(z)/\Gamma(z)$ is denoted by $\psi(z)$. This function has the properties:

$$\psi(z+1) - \psi(z) = \frac{1}{z}$$

$$\psi(1) = -C = -0.5772157\ldots$$

$$\psi(n+1) = -C + 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

$$\text{for } n = 1, 2, 3, \ldots$$

$$\psi(^1/_2) = -C - 2\ln 2$$

$$\psi(z) = -C + \sum_{n=0}^{\infty}\left(\frac{1}{1+n} - \frac{1}{z+n}\right)$$

Beta function. The beta function generalizes the binomial coefficient (N/n) to noninteger values of N and n. The beta function $B(x, y)$ is by definition equal to $\Gamma(x)\Gamma(y)/\Gamma(x+y)$. Therefore one has $(N/n)^{-1} = (N+1)B(n+1, N-n+1)$.

Some of the definite integrals, whose values can be expressed by the gamma function, the psi function, and the beta function are

$$\int_0^{\infty} \exp(-at^n)t^z\,dt$$

$$= (z+1)^{-1}a^{-(z+1)/n}\Gamma\left(\frac{z+n+1}{n}\right)$$

$$(\operatorname{Re} a > 0,\ \operatorname{Re} z > -1)$$

$$\psi(z) = \int_0^{\infty}\left(\frac{e^{-t}}{t} - \frac{e^{-zt}}{1-e^{-t}}\right)dt$$

$$= -C + \int_0^1 \frac{1-t^{z-1}}{1-t}\,dt \qquad (\operatorname{Re} z > 0)$$

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt$$

$$= 2\int_0^{\pi/2} \sin^{2x-1}\zeta\,\cos^{2y-1}\zeta\,d\zeta$$

See BESSEL FUNCTIONS.

Josef Meixner

Bibliography. L. C. Andrews, *Special Functions of Mathematics for Engineers*, 2d ed., 1998; E. Artin, *The Gamma Function*, 1964; A. Erdelyi (ed.), *Higher Transcendental Functions*, vol. 1, 1953; Z. X. Wang and D. R. Guo, *Special Functions*, 1989; E. T. Whittaker and G. N. Watson, *Course of Modern Analysis*, 4th ed., 1940, reprint 1997.

Gamma-ray astronomy

The study of gamma rays of cosmic origin. This vast spectral domain extends from an energy of 0.01 MeV (wavelength of 10^{-10} m), the adopted boundary between x-ray and gamma-ray photons, to 10^{11} MeV (10^{-23} m), an experimental barrier imposed by the extreme scarcity of ultrahigh-energy photons. *See* GAMMA RAYS.

Low-energy (or soft) gamma-ray astronomy (up to a few megaelectronvolts) deals mainly with processes in dense media, such as plasmas confined close to neutron stars and black holes. It also concerns cosmic sites where monoenergetic photons are released either by deexcitation of atomic nuclei (nuclear lines) or by positron annihilation (the 0.511-MeV line). Gamma-ray astronomy at higher energies relates to emissions induced by relativistic particles throughout the entire interstellar medium, as well as in the vicinity of some neutron stars and in the powerful jets beamed by active galactic nuclei. The penetration power of gamma-ray photons enables exploration of regions that are hidden at other wavelengths, such as the galactic center region, as well as the first stages of the universe, since the cosmos is particularly transparent to gamma rays (with the exception of ultrahigh-energy photons).

Because photons in the gamma-ray regime are completely absorbed by the Earth's atmosphere, gamma-ray detectors are placed on board high-altitude balloons, or, better still, artificial satellites. Ground-based telescopes, making use of the upper-atmospheric layers as a detector, operate successfully in the very high energy gamma-ray band. The first generation of gamma-ray telescopes included the U.S. *SAS-2* satellite launched in 1972 and the European *COS-B* satellite launched in 1975. Missions with much greater capabilities, including the French SIGMA telescope on board the Russian spacecraft *GRANAT* launched in 1989, and the large U.S. *Compton Gamma-Ray Observatory*, launched from the space shuttle in 1991, have obtained major scientific results.

At present, the gamma-ray sky in the band from 0.1 to 10 MeV is investigated by the European Space Agency's *INTEGRAL* (*International Gamma-Ray Astrophysics Laboratory*) satellite, launched in 2002. The *GLAST* (*Gamma-Ray Large-Area Space Telescope*) mission, a collaboration of the National Aeronautics and Space Administration (NASA) and the Stanford Linear Accelerator Center (SLAC), to explore the sky at higher energy, is scheduled for launch in 2007. Low-energy gamma-ray focusing experiments (*Simbol-X*, *MAX*) are foreseen in the more distant future.

Instrumentation

The information obtained by gamma-ray instruments is limited by various astronomical and instrumental factors.

Astronomical limitations. The astronomical limitations result from the weak intrinsic fluxes of gamma rays and from background problems. A typical gamma-ray observation requires long exposures, up to 1 month in duration, to obtain significant data. The main background originates when cosmic-ray particles interact with the instrument to give secondary gamma rays. Because the observations are limited by background noise, sensitivity improves only in proportion to the square root of the detector area.

Instrumental limitations. In other regions of the electromagnetic spectrum, sensitivity is increased by the straightforward method of gathering large numbers of photons and concentrating them to form an image, by means of arrangements of reflectors or lenses. Such a method has not been applied to gamma-ray telescopes, since reflection or refraction of gamma-ray photons is nearly impossible. Use of grazing-incidence mirrors involving classical reflecting surfaces would seem impracticable beyond the lower limit ascribed to the gamma-ray domain. There remains a possibility of increasing the reflection capability of metallic surfaces, for example, by means of a multilayer coating that would be able to extend the use of reflecting concentrators to energies of about 0.1 MeV. This is the principle of the US/Danish *HEFT* and the French *Simbol-X* projects. Telescopes based on Laue diffraction are also feasible, and their signal-to-noise ratio could be dramatically improved at energies up to a few megaelectronvolts. *See* X-RAY TELESCOPE.

Detection of gamma-ray photons. This involves measurements of the photon energy loss in the mass of the detector. In most cases, it is the kinetic energy imparted to charged particles by the gamma rays that is lost and measured. In photoelectric absorption, the main process at energies of less than 0.2 MeV, and in the Compton scattering process, which dominates in the 0.2–10–MeV band, part or all of the photon energy is converted to kinetic energy of secondary electrons. At higher energies, gamma rays are annihilated in the field of detector atom nuclei, producing electron-positron pairs that carry most of their incident energy. *See* COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION.

Semiconductors and crystal scintillators are widely used as gamma-ray detectors. In these materials, the secondary electrons lose energy by ionization and produce electron-hole pairs. In semiconductors, such pairs are collected to provide an electric signal; scintillation detectors require a further conversion of charged particles energy to the production of visible photons. Scintillation detectors are generally made of sodium iodide (NaI) or cesium iodide (CsI) crystals, but the energy resolution of these instruments is quite poor, making difficult the detection of spectral lines. Semiconductor germanium detectors offer very good energy resolution but require cooling to less than 100 K (−280°F). Semiconductor cadmium telluride (CdTe) or cadmium zinc telluride (CdZnTe) detectors, affording noticeable spectrometric performances combined with high detection efficiency, can operate at room temperature. Advanced position sensitive detectors, in operation in the soft gamma-ray band in the IBIS telescope on *INTEGRAL* and on NASA's *Swift* satellite (which was launched in 2004 and is devoted to the study of gamma-ray bursts), include arrays of semiconductor CdTe/CdZnTe detectors. *See* GAMMA-RAY DETECTORS; JUNCTION DETECTOR; SCINTILLATION COUNTER.

Low-energy gamma-ray telescopes. Telescopes operating in the soft gamma-ray domain, where focusing techniques have long been considered impracticable, initially used the combination of wide angle (a few degrees) collimation and on-off source chopping to yield source fluxes and location. Because of the poor angular resolution of these techniques, the firm identification of discovered sources with known astronomical objects was often possible only when the emission had a clear time signature.

Coded-aperture telescopes. A telescope for soft gamma-rays can be improved by using the coded-aperture technique to actually image celestial sources. A coded mask is a pattern of tungsten blocks that absorb gamma-ray photons and are arranged so that a given point source at infinity projects on a position-sensitive detector a pattern that is characteristic of the direction of arrival of the photons (**Fig. 1**). The position of the source in the sky is determined by comparing the observed pattern with all possible projection patterns. This technique maintains the angular resolution of a single pinhole camera, while increasing the overall effective area of the instrument. Moreover, the simultaneous measurement of the sky-plus-detector background minimizes systematic effects that result from temporal variations in the background. Both main telescopes

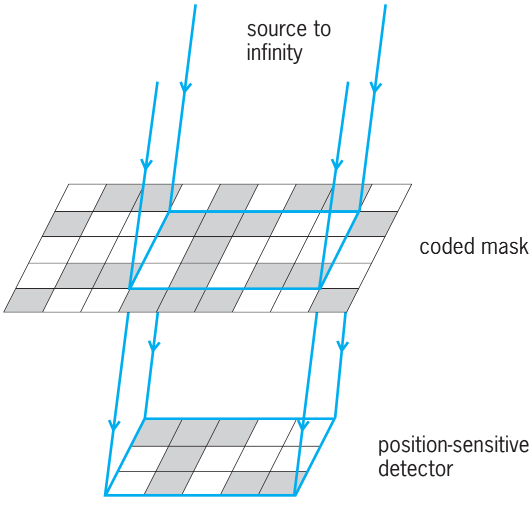


Fig. 1. Principle of a coded-mask aperture telescope whose mask is built from a 3 × 5 array extended on both axes by cyclic permutation.

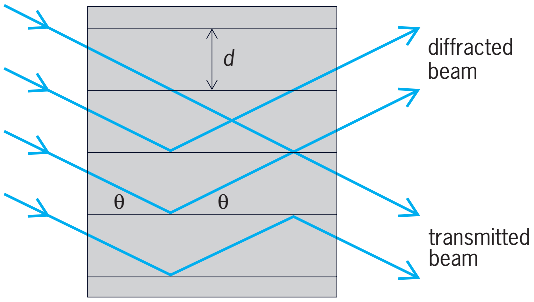


Fig. 2. Principle of Laue diffraction. A beam of very short wavelength radiation λ, entering a plane-parallel crystal at an angle θ, emerges at the same angle θ such that $\sin \theta = n\lambda/2d$, where n is an integer and d is the distance between the crystal reticular planes.

on *INTEGRAL*, IBIS and SPI, use the coded-mask principle.

Compton telescopes. Based on the detection of a Compton-scattering interaction of a gamma ray in a first detector, followed by the detection of the Compton-scattered photon in a second detector, such telescopes allow for the simultaneous determination of photon direction and energy. At megaelectronvolt energies, a set of two detectors measuring the Compton kinetics can be used in conjunction with coded-mask telescopes, in order to realize the good position resolution of coded-mask telescopes together with the low background of Compton detectors.

Concentrators. A plane-parallel crystal, intercepting a beam of very short wavelength radiation, acts as a three-dimensional diffraction array which deflects radiation by a process known as Laue diffraction (**Fig. 2**). The most advanced project of gamma-ray concentrators based on Laue diffraction, *MAX*, employs a series of concentric rings, over which are mounted several hundred germanium crystals oriented so that each deflects the incident radiation toward the same focal point. *See* X-RAY DIFFRACTION.

High-energy gamma-ray telescopes. High-energy gamma-ray observations are performed with devices derived from particle-physics detectors. High-energy gamma-ray photons interact almost exclusively via electron-positron pair production; thus that process forms the basis for the measurement principle by providing a unique signature for gamma rays and allowing a determination of the incident photon directions via the reconstruction of the electron-positron pair trajectories. After passing undetected through an anticoincidence shield sensitive to charged particles, a photon is converted to an electron-positron pair in one of the conversion foils (**Fig. 3**). Trajectories of the resulting electron and positron are measured by particle-tracking detectors. The energy deposited in an underlying calorimeter is used with the trajectory data to determine the arrival direction and total energy of the gamma rays.

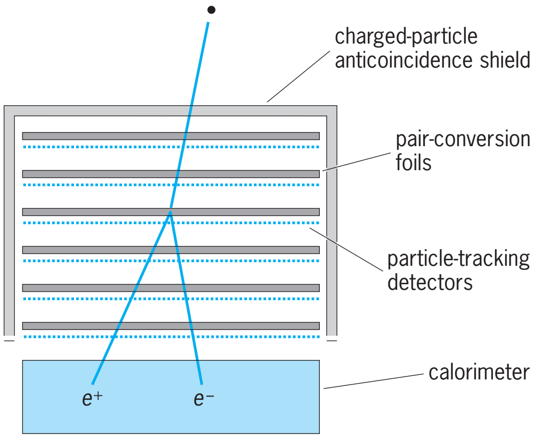


Fig. 3. Principle of a pair-conversion telescope.

Following EGRET, which operated on the *Compton/Gamma-Ray Observatory* in the 1990s, the next high-energy telescope based on these principles is the Italian pathfinder mission *AGILE* (*Astrorivelatore Gamma a Immagini L'Eggero*), to be launched in 2005, followed by the *GLAST* mission, in 2007.

Very high energy gamma-ray telescopes. As the gamma-ray energy approaches 10^5 MeV, the intensities of celestial gamma rays become too low for them to be seen with space telescopes within a reasonable observation time. However, above a few thousand megaelectronvolts, a gamma-ray photon induces in the upper atmosphere a shower of secondary relativistic particles whose propagation through the air produces a narrow beam of Cerenkov visible light that can be detected on the ground by a large parabolic mirror. Detailed studies of the Cerenkov light beam enable the determination of the arrival direction of the generating gamma ray to within 0.1° while discriminating gamma-ray-induced events from much more numerous events induced by interactions of very high energy cosmic-ray protons and nuclei. The HESS (High-Energy Stereoscopic System) telescope, for instance (**Fig. 4**), which has operated in Namibia since 2003, is a stereoscopic telescope system, where multiple telescopes view the same air shower. The initial four HESS telescopes (Phase I) are arranged in the form of a square with a side length of 120 m (400 ft) to provide multiple stereoscopic views of air showers. *See* CERENKOV RADIATION; COSMIC RAYS.

Stellar Sources

Other than the Sun, which is very near, the only stellar sources of gamma rays that are luminous enough to be observed from Earth relate to massive stars in their final stages of evolution.

Sun. Gamma-ray observations of the Sun have the objective of investigating high-energy processes that take place in the Sun's atmosphere and their relationship to the basic problems of solar activity. Solar gamma-ray emission can be detected only during solar flares. Gamma-ray continuum emissions are produced by processes, such as bremsstrahlung, involving particles accelerated during the flare. Series of sporadic gamma-ray lines are also detected. Other than the line at 2.223 MeV from deuterium formation and the line at 0.511 MeV from electron-positron annihilation, all lines are attributable to interactions

Fig. 4. HESS (High-Energy Stereoscopic System) very high energy gamma-ray telescope in Namibia.

involving energetic protons. The 2.223-MeV line implies the presence of a significant thermal neutron flux. *See* BREMSSTRAHLUNG; SUN; THERMAL NEUTRONS.

Supernovae. Major contributions to the theoretical understanding of explosive nucleosynthesis have come from the data obtained on supernova SN 1987A, which appeared on February 24, 1987, in the Large Magellanic Cloud, a nearby galaxy. Supernova explosions are considered particularly efficient in the synthesis of various radioactive nuclei. The nuclear gamma rays produced by the decay of cobalt-56 are thought to be the main power source of the supernova in the first months following the explosion. Beginning in August 1987, as the envelope of material ejected by the supernova explosion expanded and became more tenuous, the gamma-ray lines from the decay of cobalt-56 were observed (slightly preceded by the soft gamma rays of the corresponding continuum spectrum of Compton-scattered photons). In 1988, balloon-borne germanium detectors carried out fine spectroscopic studies of these lines. The evolution of the lines with time gives unprecedented constraints on explosive nucleosynthesis models and on the mixing of freshly synthesized material deep inside the supernova envelope. Gamma-ray line emissions from long-lived isotopes synthesized in supernova explosions (such as titanium-44) have also been observed in the three-century-old Cassiopeia A supernova remnant. However, a deep survey made in the galactic center region by *INTEGRAL* at the titanium-44 line energy did not reveal any sources, although ten had been expected. This discrepancy remains an unresolved issue in understanding the nucleosynthesis of titanium-44 in supernovae. *See* NUCLEOSYNTHESIS; SUPERNOVA.

Neutron stars and stellar black holes. Virtually all types of compact stars have significant high-energy emission luminosity. Many gamma-ray sources have been firmly identified with neutron stars (isolated and in binary systems) and stellar black holes.

Spin-powered neutron stars. Attested by the discovery of radio pulsars in 1967, the capability of neutron stars to accelerate relativistic electrons was confirmed a few years later by the discovery of the gamma-ray emission of the Crab and Vela pulsars. The energy spectrum of the Crab pulsar, as observed up to 10^4 MeV, suggests that the gamma radiation results from the synchrotron emission of relativistic electrons in the intense magnetic fields (of the order of 10^8 tesla or 10^{12} gauss) which prevail in the close vicinity of newly formed neutron stars. It is generally agreed that the ultimate source of the radiated energy is the rotational energy of the neutron star. Electrons are accelerated to very high energies by the huge electric fields induced by the pulsar rotation, as implied by the gamma-ray emission detected above 5×10^5 MeV in the direction of the Crab Nebula by atmospheric Cerenkov devices. It has been suggested that such a very high energy gamma-ray emission results from inverse Compton processes where ultrarelativistic electrons accelerated by the Crab pulsar energize soft photons within the Crab nebula. *See* CRAB NEBULA; SYNCHROTRON RADIATION.

Nevertheless, much of the basic workings of pulsars remains unknown: How and where does the particle acceleration take place? What is the shape of the particle beam, and how is its energy converted to photons? High-energy gamma-ray studies provide an essential tool for answering these questions. The stage has been set by the seven pulsars detected by the *Compton Gamma-Ray Observatory*, which are being studied in more detail by *INTEGRAL* (**Fig. 5**). These include the brightest persistent sources in the 10^2–10^4-MeV band: the Crab and Vela pulsars and Geminga, which was also found in 1992 to be a gamma-ray pulsar. Geminga has been identified with a very weak visible star whose measured parallax indicates that it is one of the nearest known neutron stars, 450 light-years (4.3×10^{15} km or

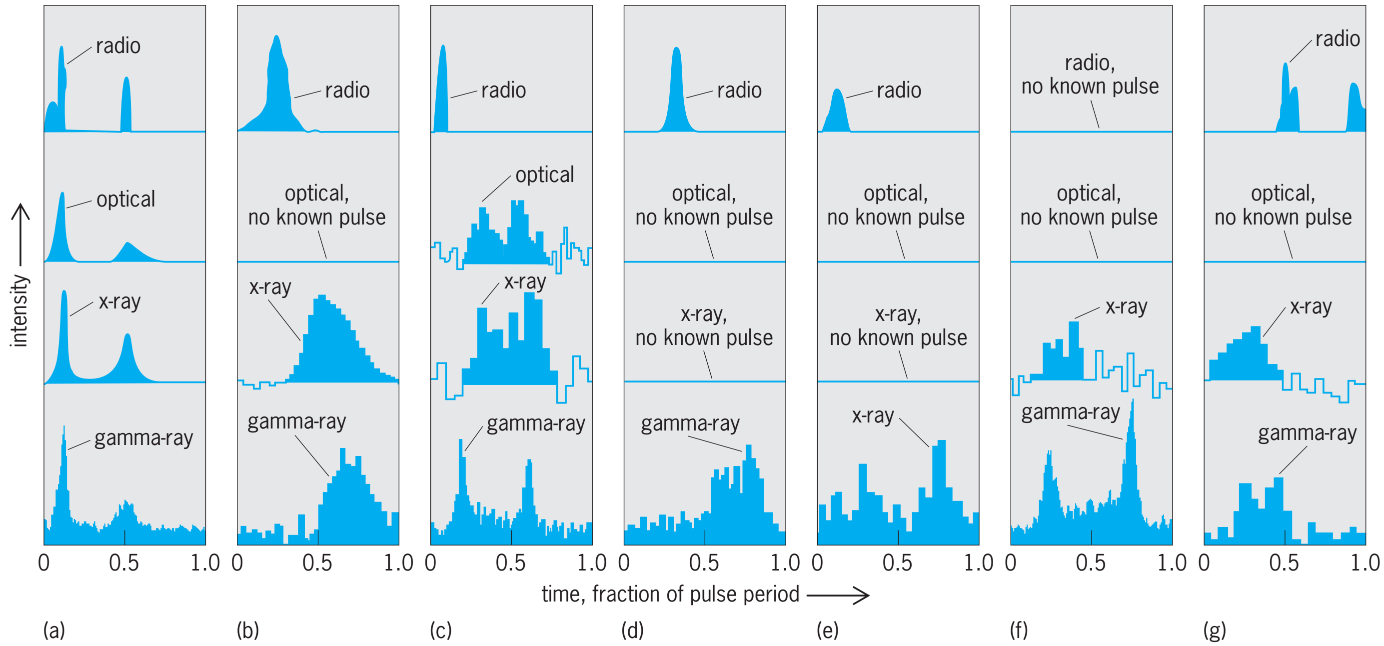


Fig. 5. Light curves of seven isolated pulsars detected in the gamma-ray range by the telescope EGRET onboard the *Compton Gamma-Ray Observatory*. Light curves are shown in the radio band, the optical band, x-rays, and the gamma-ray energy domain. (*a*) Crab pulsar, period ~33 ms. (*b*) PSR 1509-58, period ~150 ms. (*c*) Vela pulsar, period ~89 ms. (*d*) PSR 1706-44, period ~102 ms. (*e*) PSR B1951+32, period ~39 ms. (*f*) Geminga, period ~237 ms. (*g*) PSR 1055-52, period ~197 ms. (*Courtesy D. Thomson, NASA/GSFC*)

2.6×10^{15} mi) from the Earth. The Crab and Vela pulsars have also been detected at very high energy, 10^5–10^6 MeV, by several Cerenkov experiments (including Whipple, CAT, and CANGAROO). *See* GEMINGA; NEUTRON STAR; PULSAR.

Accreting compact stars. The compact remnant of an exploded star belonging to a close binary system exerts such an attraction on the external layers of its companion star that matter transfers may occur. This transfer results in a thick plasma disk, the accretion disk, which encircles the collapsed star, itself enshrouded in a dim corona. Heated by viscous stresses, the inner part of the disk releases intense thermal radiation in the x-ray band. Gamma rays could be produced by the Compton upscatter process, in which hot (fast moving) electrons near the compact object scatter and kick the abundant x-ray photons to higher energies. The electrons, which gain energy while falling toward the compact object, lose energy during this scattering process, so that an equilibrium temperature is obtained close to the compact object surface. This temperature is particularly high if the collapsed star is a black hole, and hence lacks a solid surface that could reemit a lot of soft x-ray photons, which are particularly efficient in cooling the electron population.

Almost all known accreting black-hole systems produce strong and variable gamma-ray emission. In systems with low-mass (less than one solar mass) companion stars, huge outbursts called x-ray novae are observed, such as Nova Muscae 1991, which for one week was the brightest source in the soft gamma-ray sky. On January 20–21, 1991, it underwent an eruptive episode that appeared in the form of a spectral feature centered around 0.5 MeV (**Fig. 6**), probably related to electron-positron annihilation processes.

Several accreting black holes have been observed in the central region of the Milky Way Galaxy (**Fig. 7**), while the galactic nucleus itself has been detected at a surprisingly weak level in the soft gamma-ray regime by *INTEGRAL*. This weakness may be due to the different type of accretion occurring close to the supermassive black hole sleeping in the galactic center. Accurate radio observation of the source 1E 1740.7-2942, located only 50 arc-min from the nucleus, and of a few similar black-hole candidates, has revealed that accreting stellar black holes could

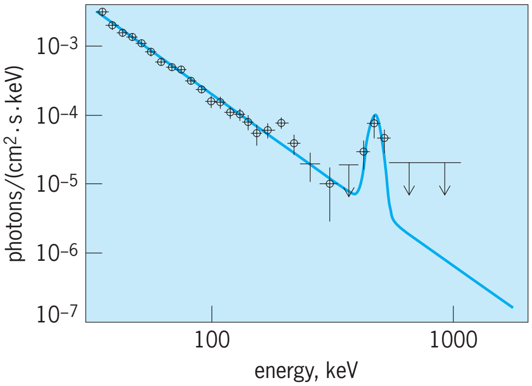


Fig. 6. Energy spectrum of Nova Muscae 1991 as obtained by the SIGMA telescope on board the *GRANAT* satellite on January 20–21, 1991 (crosses). The most straightforward model (solid line) consists of superposing a gaussian-profile line centered at 0.480 MeV onto a power-law spectrum of spectral index −2.4. (*After A. Goldwurm et al., SIGMA observations of the x-ray nova in Musca, Astron. Astrophys. Suppl., 97:293–297, 1993*)

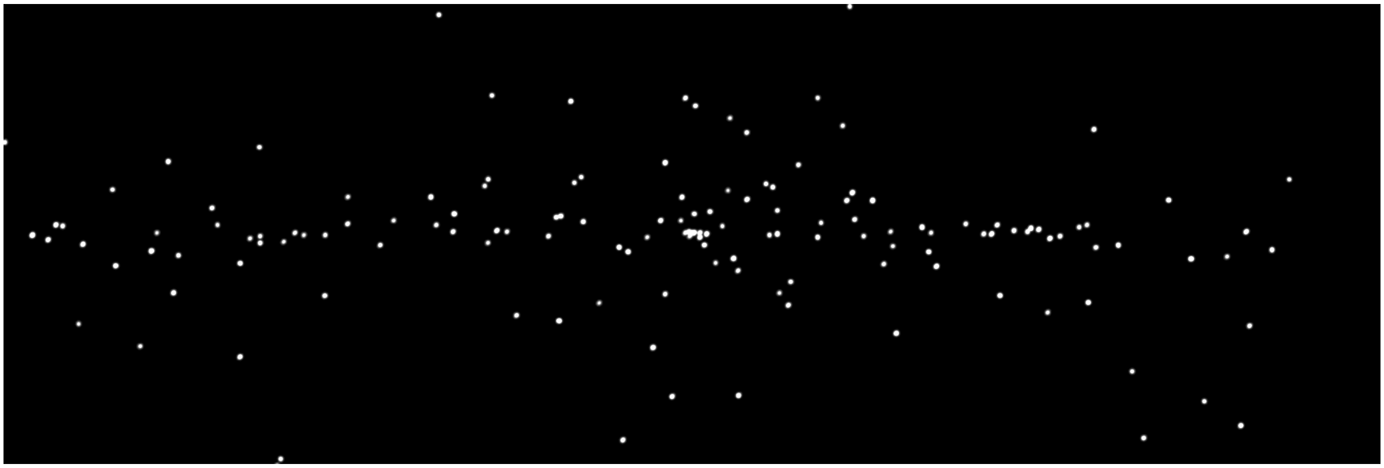


Fig. 7. Image in the 0.020–0.060-MeV band of a 90° × 30° region around the galactic center as seen by the IBIS telescope on board the *INTEGRAL* satellite. The map is in galactic coordinates; the galactic plane lies along the horizontal line at galactic latitude = 0°. More than 90 sources can be seen on this map, some of them being black holes. (*F. Lebrun et al., Compact sources as the origin of the soft γ-ray emission of the Milky Way, Nature, 428:293–296, 2004*)

generate powerful bipolar jets of relativistic particles. Featuring two basic properties of quasars (black holes and radio jets), such sources are sometimes referred to as microquasars. *See* ASTROPHYSICS, HIGH-ENERGY; BLACK HOLE; MILKY WAY GALAXY.

Other Galactic Sources

A large fraction of the cosmic gamma-ray photons originates in interstellar sites, occasionally so widely spread out that they are well resolved by gamma-ray devices, in spite of their limited angular resolution.

Cosmic-ray-induced interstellar emission. This emission results mostly from the interaction of cosmic rays (electrons, and protons) with the interstellar gas.

Relativistic electrons (of energies above a few megaelectronvolts) radiate gamma rays via the bremsstrahlung process when they are deviated by interstellar nuclei; fast protons (of energies above a few hundred megaelectronvolts) indirectly generate gamma rays when they collide with protons at rest, through the creation of π° particles (neutral pions), which decay almost instantaneously into two 68-MeV gamma rays. *See* ELEMENTARY PARTICLES.

The high-energy gamma-ray sky is dominated by radiation from the galactic plane (**Fig. 8**). Its spatial distribution and intensity can be reliably modeled from knowledge of the interstellar gas distribution. An important advance in understanding cosmic rays was the detection by the *Compton Gamma-Ray Observatory* of the broad spectral feature at 68 MeV, due to π° decay. This feature is the signature of the presence of accelerated particles throughout the galactic disk. *See* INTERSTELLAR MATTER.

Localized sources. Pointlike sources have been systematically searched for on top of the large-scale emission. Above a few megaelectronvolts, this search is particularly difficult as the telescopes operating in this energy range have a poor localization power. Indeed, the third EGRET catalog of 271 sources at energies above 100 MeV includes 170 specimens that have no established counterparts at other wavelengths. On the other hand, a number of supernova remnants are reported to be coincident with EGRET sources. These high-energy gamma-ray emissions may result from the interaction of particles accelerated in the shocks produced by collisions of the supernova remnant with the surrounding dense interstellar clouds. Although these results are somewhat ambiguous, recent observations of the remnant of SN 1006 (a supernova seen in the year 1006) performed in the very high energy band by ground-based Cerenkov telescopes demonstrate that at least electrons are accelerated to very high energies in the shocks of supernova remnants.

A large fraction of the sources observed at medium galactic latitude may be related to the local interstellar medium, and more specifically to the giant cloud complexes of Gould's Belt. In addition to the pointlike sources, two positive identifications of larger-scale features are a concentration of clouds near the star ρ Ophiuchi and the Orion cloud complex. In the

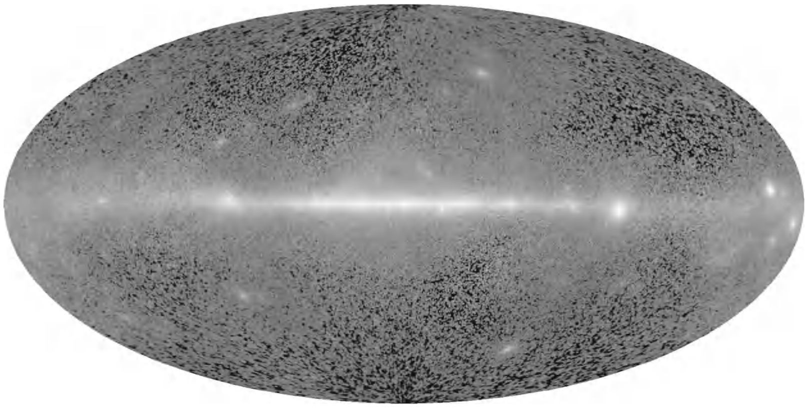


Fig. 8. Intensity map of the gamma-ray sky at energies greater than 100 MeV as measured by the EGRET experiment on the *Compton Gamma-Ray Observatory*. The map is in galactic coordinates, with the galactic plane along the abscissa. The broad band along the galactic plane is interstellar emission from cosmic-ray interactions in the Milky Way. The brightest sources close to the plane are gamma-ray pulsars. The bright sources away from the plane are active galactic nuclei or yet unidentified. (*Courtesy of the EGRET team*)

constellation Aquila, the detection of high-energy gamma-ray emission led to the discovery of interstellar clouds, later confirmed by radio observations of the carbon monoxide molecule. *See* ORION NEBULA.

Interstellar line emissions. Spectroscopic observations of the whole galactic center region, especially by the spectrometer on board *INTEGRAL*, SPI (**Fig. 9**), have demonstrated the presence of a large-scale component of 0.511-MeV radiation due to the annihilation of positrons in the interstellar medium. Such interstellar positrons were thought to be produced through the β^+ decay of radioactive nuclides produced by novae, red giants, Wolf-Rayet stars, and supernovae. However, the SPI observations imply large numbers of produced positrons, which seem inconsistent with the standard models. Alternative explanations, involving more exotic objects, such as hypernovae or dark-matter particles, may be implied by the SPI measurements. *See* POSITRON.

The discovery and confirmation of 1.809-MeV-line emission from the galactic plane (**Fig. 10**) was the first direct detection of a cosmic radioactive nucleus (aluminum-26) from the characteristic signature of its decay. As the half-life of aluminum-26 (716,000 years) is very short compared to the time scale of galactic chemical evolution, the 1.809-MeV-line observations clearly demonstrate that nucleosynthesis is currently taking place in the Milky Way Galaxy. The celestial distribution of the 1.809-MeV-line emission suggests that advanced stages of massive stars, such as Wolf-Rayet stars and supernovae, are the major sources of aluminium-26. *See* RADIOACTIVITY; WOLF-RAYET STAR.

Extragalactic Gamma-Ray Sources

With the exception of the Large Magellanic Cloud, which harbors both SN 1987A and a diffusely distributed source of high-energy gamma rays, similar to that related to the whole Milky Way Galaxy, and with the exception of the gamma-ray bursts, all localized extragalactic sources of gamma rays are active galactic nuclei, among the most energetic and distant objects in the universe. *See* GAMMA-RAY BURSTS.

Active galactic nuclei. These include radio sources such as Seyfert galaxies and quasars, all with different properties depending on the observing wavelength. The ultimate source of active galactic nuclei activity is believed to be massive (10^6–10^9-solar-mass) black holes, accreting 10–100 solar masses per year to account for their overall luminosity. An accretion disk with a collimated perpendicular jet is the favored model for explaining the luminous, broadband radiation emitted from the central engines of active galactic nuclei. While ultraviolet, x-ray, and soft gamma-ray photons are emitted somewhat isotropically from the accretion disk (several Seyfert galaxies have been detected in the soft gamma-ray band), photons at higher energies are produced by radiation mechanisms that focus gamma rays in the jet direction.

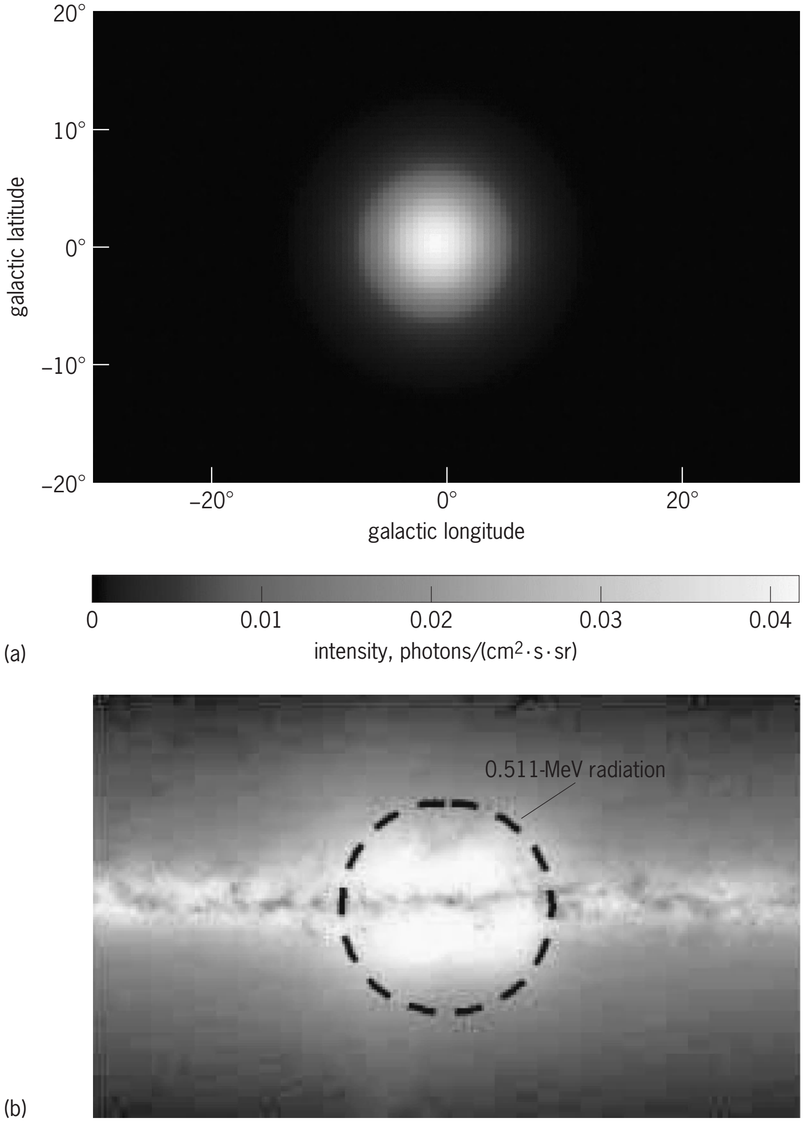


Fig. 9. Gamma-ray line emission at 0.511 MeV from electron-positron annihilation from the direction of the galactic center. (*a*) Intensity map of 0.511-MeV emission as measured by the SPI telescope on the *INTEGRAL* satellite. The map is in galactic coordinates, with the galactic plane along the abscissa. (*From G. Weidenspointner et al., SPI observations of positron annihilation radiation from the 4th galactic quadrant: Sky distribution, in Proceedings of the 5th INTEGRAL Workshop on The INTEGRAL Universe, Munich, pp. 133–139, ESA SP-552, 2004*) (*b*) Distribution of 0.511-MeV radiation superposed on an infrared view of the Milky Way, showing that the distribution coincides with the galactic bulge. (*CESR/Toulouse*)

Only BL Lacertae objects as well as highly variable and strongly polarized quasars (collectively referred as blazars) have been detected at high energies, because blazar jets are favorably aligned along the line of sight. A total of 66 high-confidence identifications of blazars are indexed in the third EGRET catalog, including very distant objects (up to redshift $z = 2.3$). In addition, 27 lower-confidence potential blazar identifications are noted. A few nearby blazars have been detected at very high energies by ground-based Cerenkov telescopes. Their spectra extending to energies above 10^6 MeV, variability time scales measuring less than a day, and apparent luminosities often exceeding 10^{15} solar luminosity are evidence that extremely energetic sites of efficient relativistic particle acceleration must be involved. High-energy gamma rays are thought to result from Compton scattering of soft photons by energetic nonthermal electrons in the jet. The source of the soft photons is controversial: synchrotron self-Compton models, employing

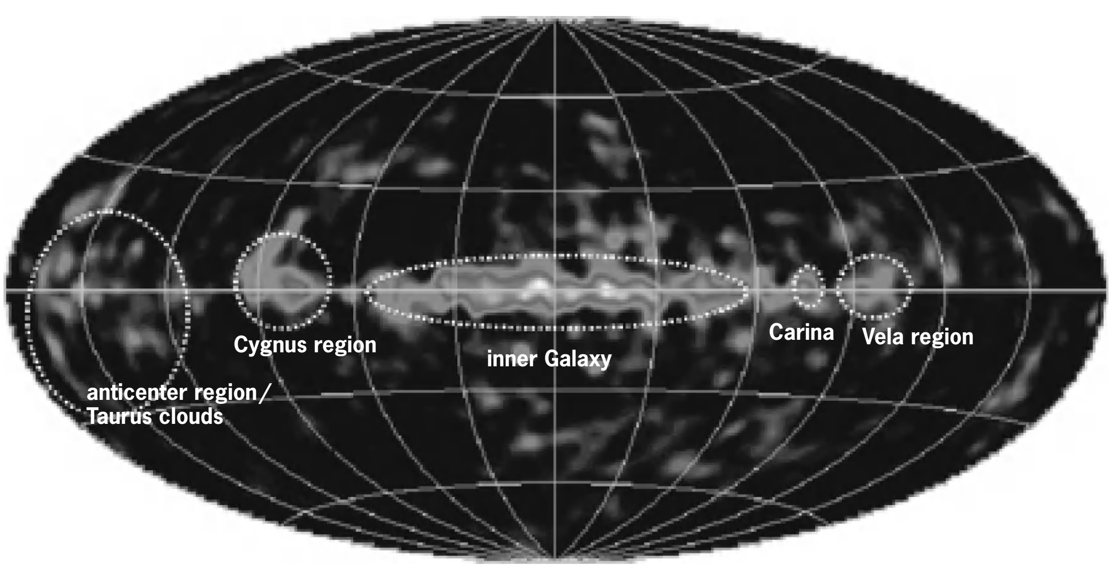


Fig. 10. Intensity map of the gamma-ray line emission at 1.809 MeV from aluminum-26 decay as measured by the COMPTEL experiment on the *Compton Gamma-Ray Observatory.* The map is in galactic coordinates, with the galactic plane along the abscissa. (*After U. Oberlack, Ph.D. thesis, Technische Universität München, 1997*)

the electrons' own synchrotron radiation, are one possibility. *See* GALAXY, EXTERNAL; QUASAR.

Cosmic background radiation. An apparently isotropic, presumably extragalactic, radiation has been measured in the gamma-ray regime up to about 150 MeV. The whole cosmic gamma-ray background could be explained in term of composite light from a large number of faint active galactic nuclei such as Seyfert galaxies at low energy and blazars at high energy. Other more exotic possibilities are relics from some yet-unknown high-energy process in the early universe.

Philippe Laurent

Bibliography. N. Gehrels and J. Paul, The new gamma-ray astronomy, *Physics Today*, pp. 26–32, February 1998; R. C. Hartman et al., The Third EGRET Catalog of High-Energy Gamma-Ray Sources, *Astrophys. J. Suppl.*, 123:79–202, 1999; Special Letters Issue: First science with *INTEGRAL, Astron. Astrophys.*, 411:L1–L460, 2003.

Gamma-ray bursts

Short cosmic blasts of very high energy electromagnetic radiation (gamma radiation) that are recorded at an average of about once per day by detectors placed above the Earth's atmosphere. At distances of billions of light-years, the energy emitted in a gamma-ray burst (GRB) is more than a billion billion (10^{18}) times the energy emitted each second by the Sun. The mechanism for the origin of gamma-ray bursts remains a focus of current research. Gamma-ray bursts longer than 2 seconds appear to be associated with the supernova explosions of massive stars in distant galaxies. Recent evidence suggests that shorter gamma-ray bursts are the result of merging neutron stars or black holes. There are indications that both types of gamma-ray bursts lead to the birth of black holes. *See* GAMMA RAYS.

History of observations. In 1963, the United States and other nations signed the Nuclear Test Ban Treaty, which forbade the testing of nuclear weapons in or above the atmosphere. In order to ensure compliance with this treaty, the United States launched the *Vela* series of Earth-orbiting satellites to look for violations which would produce blasts of gamma radiation. No violations of the Nuclear Test Ban Treaty were ever reported. However, in 1973 Ray W. Klebesadel, Ian B. Strong, and Roy A. Olson announced the discovery of the first gamma-ray bursts and provided evidence of their cosmic origin based on the data from the *Vela* satellites.

During the next 20 years, an international array of satellites slowly built up a catalog of information about the brightness and energies of the gamma rays in scores of gamma-ray bursts. Some of these satellites were located in the solar system at great distances from Earth, while others were in near Earth orbits. The combined data from this Interplanetary Network (IPN) allowed the positions of some of the bursts to be located with greater precision. However, until the notable event that occurred on March 5, 1979, the identification of the gamma-ray bursts with familiar objects that radiate at, for example, visible wavelengths of light, proved elusive.

March 1979 event. On March 5, 1979, at least nine satellites in the Interplanetary Network were swamped with a blast of gamma radiation, 100 times stronger than any burst detected previously. Because of the large number of detections converging from different angles, the Interplanetary Network was able to precisely locate the burst. Surprisingly, it emanated from a supernova remnant in our neighboring galaxy, the Large Magellanic Cloud, at a distance of about 150,000 light-years. A sinusoidal signal was seen in the gamma-ray light as the intensity decayed, followed hours later by another weaker burst, with many more bursts occurring over the next month. Studies of gamma-ray bursts for the next decade assumed that the March 5, 1979, event was a prototype for all gamma-ray bursts, and many scientists searched for sinusoidal oscillations and repeated weaker bursts at the positions of seemingly similar events. However, it is now known that the March 5, 1979, event was not a "classic gamma-ray burst" but the defining member of a new class of cosmic explosions—the soft Gamma repeaters (SGRs). These seemingly similar detonations result from starquakes on the surfaces of rotating neutron stars which possess extremely strong magnetic fields, with strengths up to 10^{11} tesla. *See* MAGELLANIC CLOUDS; NEUTRON STAR.

CGRO observations. The *Compton Gamma-Ray Observatory* (*CGRO*), launched into Earth orbit by the National Aeronautics and Space Administration (NASA) on April 5, 1991, had a complement of four gamma-ray-detecting instruments, including one designed specifically to study gamma-ray bursts: the Burst And Transient Source Experiment (BATSE). During its 9-year lifetime, BATSE detected and located over 2700 gamma-ray bursts; **Figure 1** shows their positions and fluxes on a map of the sky. There is no obvious clustering of the bursts near the plane of the Milky Way Galaxy (horizontally across the center of the figure) or anywhere else. This result surprised almost all astronomers, as it had been widely believed that the gamma-ray bursts must arise from within the Milky Way Galaxy.

Two possible interpretations were advanced to explain the lack of any pattern in the gamma-ray burst skymap: either the bursts were very close to the solar system, or they were very far away, at cosmic distances. In this latter case, however, the energies required to power the bursts were so large that scientists had trouble explaining the cause of the bursts. The evidence that gamma-ray bursts were truly located across the universe did not appear until 1997.

Discovery of afterglows and distances. On February 28, 1997, a gamma-ray burst was observed by detectors on board *BeppoSAX*, an Italian-Dutch satellite. Eight hours later, after maneuvering the satellite to use a different set of detectors, the *BeppoSAX* observing team discovered an x-ray source that was located at the position of the initial blast of gamma-rays. A second observation 3 days later showed that the source was fading—they had discovered the first x-ray "afterglow" of a gamma-ray burst, radiation emitted by the cooling embers of the much hotter initial explosion. The *BeppoSAX* team then alerted observers at ground-based telescopes, leading to the discovery of the first visible-light afterglow. On May 8, 1997, the *BeppoSAX* team located another x-ray afterglow, and quick work resulted in the first determination of a distance to a gamma-ray burst—about 7×10^9 light-years—providing the first real evidence that gamma-ray bursts are located at cosmic distances. *See* X-RAY ASTRONOMY.

The distances to the gamma-ray bursts cannot be obtained solely by using gamma-ray data, but result from the measurement of red-shifted spectral lines in the visible-light spectrum of the afterglow. (Spectral lines are redshifted due to the expansion of the universe. As the universe expands, light waves traveling great distances are stretched by the expansion. Longer-wavelength light is "redder," hence the term redshift. With a few additional assumptions about the expansion history of the universe, a measured redshift can then be converted into a distance.) As of 2006, distances had been measured to over 70 gamma-ray bursts: the farthest, detected on September 4, 2005, by NASA's *Swift* satellite, is at a distance of about 12.8×10^9 light-years (redshift of 6.3), and the average distance is 9.8 billion light-years (redshift of 1.7). For comparison, the most distant galaxy seen so far is a quasar at about 13×10^9 light-years (or a redshift of 6.5). Evidence has accumulated that bursts with afterglows are located in regions of distant galaxies that are actively forming stars. *See* COSMOLOGY; GALAXY, EXTERNAL; QUASAR; REDSHIFT.

Observations with burst alert systems. As BATSE accumulated observations of gamma-ray bursts, ground-based observers became eager to help unravel the mystery of the bursts' origins. To facilitate these efforts, in 1993 Scott Barthelmy created a system for distributing information about the bursts, including crude positions, to interested observers. Originally called BACODINE (BAtse COordinates DIstribution NEtwork), the system used telephones, pagers, and eventually the Internet to notify interested astronomers of burst arrivals within seconds. In 1998, the system was renamed the Gamma-ray Burst Coordinates Network (GCN). It now sends out burst alert notices from several satellites, and also provides a repository for the ground- and satellite-based observers to post their scientific results. In many cases the burst alerts are directly transmitted into the control systems for robotic telescopes linked through the Internet. The telescopes then autonomously repoint to observe the sky locations of the bursts.

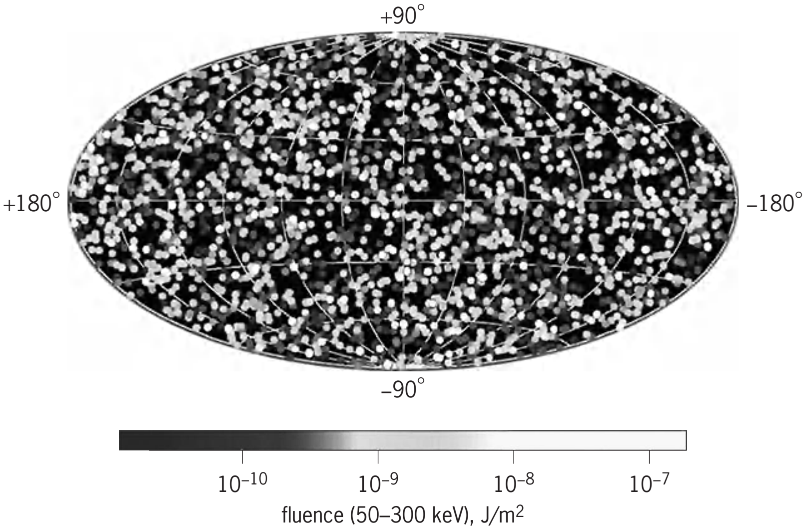


Fig. 1. Equal-area projection map of the locations of over 2700 gamma-ray bursts on the sky as measured by the BATSE detector on the *Compton Gamma-Ray Observatory*. Different shades indicate the total energy emitted (fluence) for each burst. The plane of the Milky Way Galaxy spans the Equator. (*NASA/CGRO/BATSE Team*)

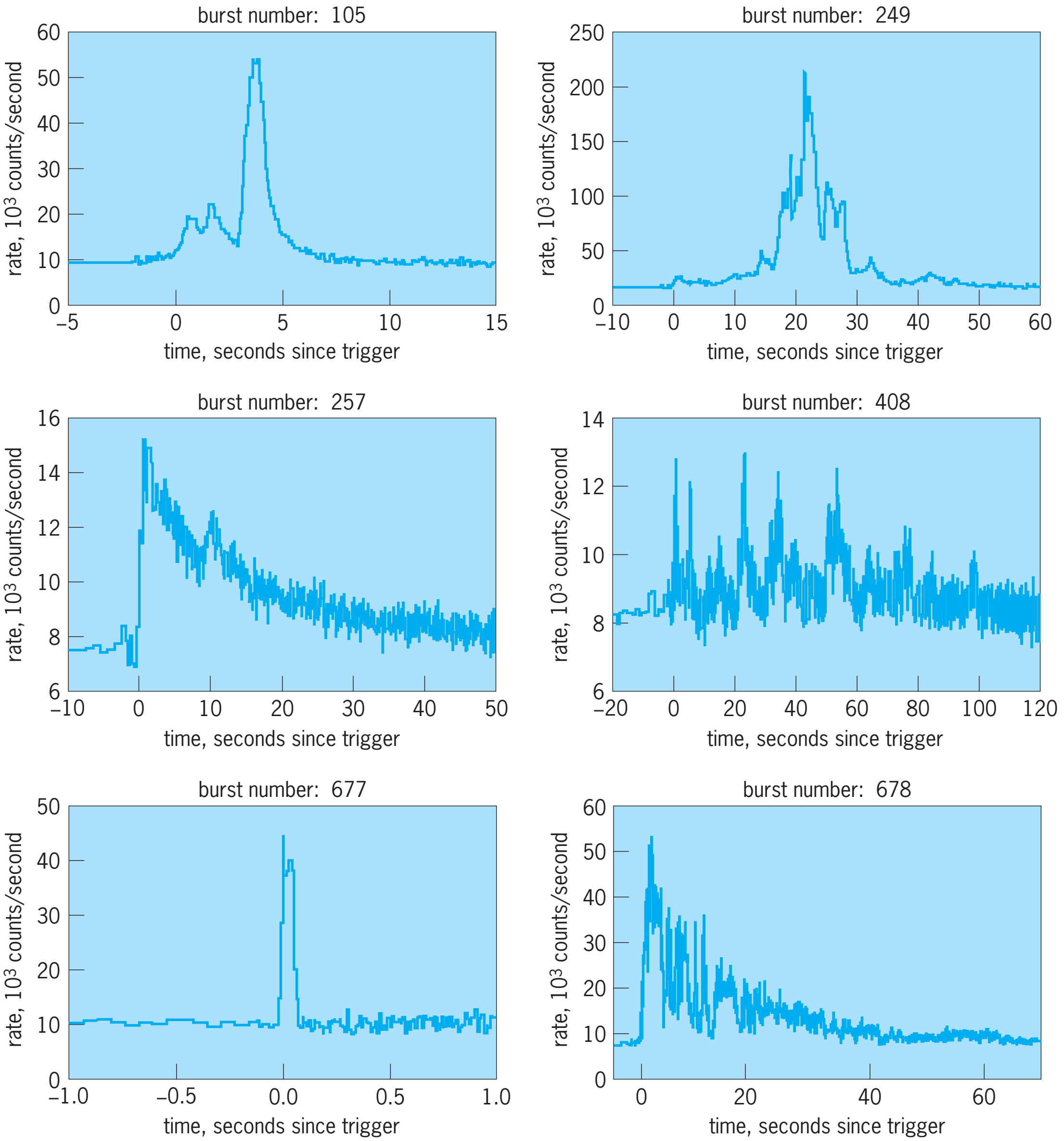


Fig. 2. Time histories of various gamma-ray bursts detected by BATSE on the *Compton Gamma-Ray Observatory*, showing a broad range of time scales. (*NASA/CGRO/BATSE Team*)

The Robotic Optical Transient Search Experiment (ROTSE) was the first autonomous telescope to be connected to the Internet-driven burst alert system. After over 5 years of monitoring the burst alert notices, on January 23, 1999, ROTSE quickly repointed to acquire a series of images beginning 22 seconds after the initiation of a burst reported by BATSE. It was the first burst "caught in the act." ROTSE's images of the rapidly fading "prompt" optical emission (rather than the afterglow) were so bright (9th magnitude) that the optical flash could have been seen by an observer using binoculars. It was realized that, with a sufficiently quick response, it would be possible to observe optical light that was emitted during the burst itself, as well as the afterglows. Following the launch in November 2004 of NASA's *Swift* satellite, the hunt for the origin of gamma-ray bursts intensified as many additional ground-based observers joined the burst chase. One of the first *Swift* gamma-ray bursts (on January 17, 2005) was observed using onboard x-ray, ultraviolet, and optical telescopes while flaring activity continued for hundreds of seconds. *Swift*'s rapid response has allowed many ground-based robotic telescopes to study bursts in their early phases, as well as the compiling of detailed observations of afterglow emission across the electromagnetic spectrum.

Observational properties. The time histories of the brightnesses of gamma-ray bursts show a remarkable range of variation. (**Fig. 2**). Some bursts last only a few milliseconds, while others show multiple peaks that last for minutes. These widely varying characteristics pose a difficult challenge to scientists seeking to understand the origin of the bursts.

There are two classes of gamma-ray bursts that are primarily distinguishable by their durations and spectral characteristics: short bursts, with durations less than 2 seconds and more energetic (harder) spectra; and bursts longer than 2 seconds, with less energetic (softer) spectra. Almost all the afterglow measurements that have been obtained so far are from the longer, softer class of gamma-ray bursts. To date, there are only a handful of short, hard bursts with observed afterglows: on May 9, 2005, NASA's *Swift* satellite detected the first x-ray afterglow, and on July 7, NASA's *HETE-2* satellite detected a short burst that led to the first optical afterglow, seen in data collected with ground-based telescopes.

Possible origins. With the knowledge that the longer-duration gamma-ray bursts are being created at vast distances, in star-forming regions of galaxies, speculation has increased that they are somehow related to supernova explosions of massive stars that have reached the end of their relatively short, nuclear-fueled lives. Such massive stars preferentially form in denser regions in galaxies where there is an abundance of material to seed the star-formation process, and burn very brightly, exhausting all their fuel within a few million years. When a massive star exhausts its nuclear fuel, it will collapse inwardly while flinging out its outer layers in a supernova explosion. If a large piece, greater than 3 solar masses (6×10^{30} kg or 1.3×10^{31} lb), of the inner core remains intact, it can collapse directly into a black hole. *See* BLACK HOLE; STELLAR EVOLUTION; SUPERNOVA.

There is direct evidence that two gamma-ray bursts are associated with supernovae. On March 29, 2003, an extremely bright gamma-ray burst was detected by NASA's *HETE-2* satellite, and its position was relayed by the Gamma-ray Burst Coordinates Network. This well-studied burst produced an afterglow that was detected by visible-light observers at ground-based telescopes around the world. A detailed study of the afterglow light revealed the characteristic signature of a supernova, located only 2×10^9 light-years away—one of the closest gamma-ray bursts ever detected. Further studies of this supernova proved conclusively that it was related to the gamma-ray burst, providing indisputable proof that the two events were connected. On February 18, 2006, *Swift* detected a gamma-ray burst precursor to a supernova, which occurred later that week, as astronomers worldwide trained their telescopes on the amazing event. This burst, at a distance of 440 million light-years, was the second-closest burst on record, offering another clue to the gamma-ray burst–supernova connection.

In the supernova scenario—named the "collapsar" or "hypernova" model—an extremely massive star, greater than 100 times the mass of the Sun (2×10^{32} kg or 4×10^{32} lb), explodes, and then its inner core collapses to form a black hole. The newly born black hole then accretes some of the swirling and collapsing matter to form jets. These jets carry matter and energy outward, plowing through the collapsing layers of matter, and creating internal shocks that produce gamma rays (**Fig. 3**). The interaction of the jets with the collapsing material produces a wide variety of different gamma-ray time histories, forming gamma-ray bursts with different numbers of peaks and durations—depending on the amount and distribution of the matter that is being shocked. When the jets impact the more tenuous matter blown off earlier by the star as it evolves, additional shocks create x-rays, and then visible and finally radio light as the matter cools. These lower-energy emissions are seen as the afterglow, which varies depending on the density and temperature of the ambient material in the nearby vicinity. Radio observations, at the Very Large Array have shown that the radio emissions from gamma-ray bursts are highly collimated into beams or jets, thus reducing the energy needed to power the bursts. (Since the gamma rays that we see are in narrow jets aimed in our direction, rather than emitted uniformly in all directions, the energy inferred in the gamma-ray burst is lowered by a factor of around 100.) *See* RADIO TELESCOPE.

Fig. 3. Conception of a hypernova explosion with gamma-ray emitting jets. (*Dave Ambrecht, General Dynamics*)

Important clues to the origin of the shorter gamma-ray bursts—those with duration less than 2 s—were revealed in 2005 with the discovery of afterglows by *Swift* and *HETE-2*. These observations are consistent with the orbital decay of two compact objects in a binary system: two neutron stars, black holes, or perhaps one of each. Einstein's theory of gravity (general relativity) predicts that binary orbits will decay by the emission of gravitational radiation. Several binary pulsar (neutron star) systems are now known (including the original system, discovered by Russell Hulse and Joseph Taylor). As the two objects finally gravitate together, their death spiral will lead to the formation of a black hole, accompanied by a relatively small disk of debris. The formation of jets from the accretion of this disk onto the rapidly spinning black hole can release a tremendous surge of gamma rays (**Fig. 4**). In this case, there is little nearby material to create afterglow emission, and the gamma rays can escape the stellar envelope more quickly, as not much obscuring matter is present. Hence the gamma-ray bursts that are formed in

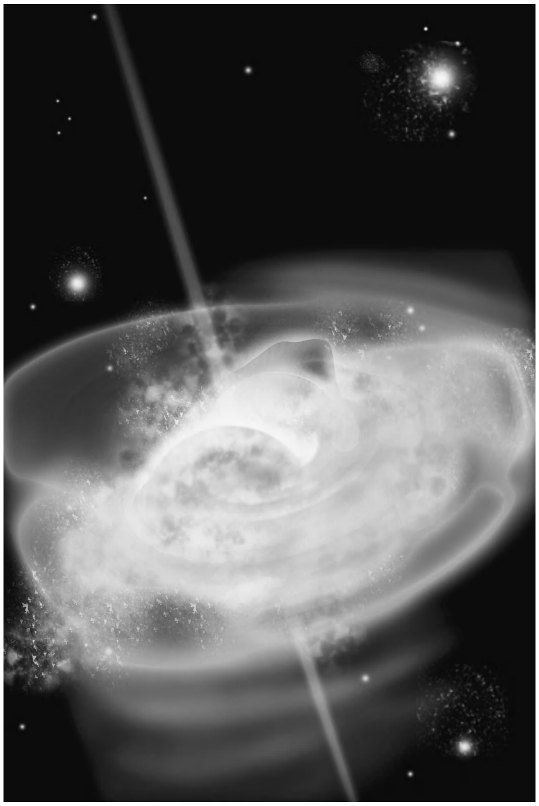


Fig. 4. Conception of merging neutron stars. (*Aurore Simonnet, Sonoma State University*)

the merger scenario are much shorter, and should not be accompanied by long-lived afterglows. The x-ray and optical afterglows from the short bursts seen with *Swift* and *HETE-2* have confirmed this picture. The afterglows are short-lived, with no evidence for any supernovae. In addition, the locations in which bursts occurred are older, elliptical galaxies, or galactic regions that are not actively forming stars. In both the collapsar and merger scenarios, accounting for the longer and shorter gamma-ray bursts respectively, the bursts appear to signal the creation of black holes. *See* GRAVITATIONAL RADIATION; PULSAR; RELATIVITY.

Present and future studies. In mid-2006, eight satellites were observing gamma-ray bursts, and participating in the Interplanetary Network: NASA's *HETE-2*, *RHESSI*, *Mars Odyssey*, and *Swift* spacecraft; the *Ulysses* mission of NASA and the European Space Agency (ESA); NASA's *Wind* experiment on the Russion *Konus* mission; the Japanese *Suzaku* satellite; and ESA's *INTEGRAL* mission. Future missions that are planned to study gamma-ray bursts include the *Gamma-ray Large Area Space Telescope* (*GLAST*), scheduled for launch in 2007.

Lynn Cominsky

Bibliography. N. Gehrels et al., The *Swift* gamma-ray burst mission, *Astrophys. J.*, 611:1005–1020, 2004; N. Gehrels, L. Piro, and P. J. Leonard, The brightest explosions in the universe, *Sci. Amer.*, 287(6):84–92, December 2002; C. Kouveliotou, R. C. Duncan, and C. Thompson, Magnetars, *Sci. Amer.*, 288(2):34, 2003; F. Reddy, The star-splitting jets of supernovae, *Astronomy*, 32(10):44, October 2004; D. Reichart, The gamma-ray burst/supernova connection, *Mercury*, 32(5):15, September/October 2003; G. Schilling, Catching gamma-ray bursts on the wing, *Sky & Telesc.*, 107(3):32–40, March 2004; G. Schilling, *Flash! The Hunt for the Biggest Explosions in the Universe*, Cambridge University Press, 2002; J. Wise, How to outsmart a gamma-ray burst, *Pop. Sci.*, October 2004.

Gamma-ray detectors

Instruments that register the presence of gamma (γ) radiation. Such detectors convert some or all of the energy of gamma radiation into an electrical signal. Most instruments are capable of detecting individual gamma-ray photons (or quanta), each of which produces a short (0.1–5-microsecond) current pulse at the detector output. The output pulses may be made visible on an oscilloscope, made audible through a speaker (such as the familiar Geiger counter), or be electronically processed further, depending on the application. *See* GAMMA RAYS; OSCILLOSCOPE.

Gamma-ray detectors range from hand-held devices capable of giving some indication of the intensity of a radiation field, to devices that accurately measure the energy and event time of individual photons reaching detectors assembled into a single complex instrument. These diverse detectors are widely used in industry, medicine, and research.

Response to gamma radiation. In common with most radiation detectors, gamma-ray detectors respond not to the radiation but to its secondary effects, in this case energetic electrons. Photons have neither mass nor charge and pass through matter rather easily. In so doing, they lose energy by (1) elastic scattering of electrons (Compton effect), (2) electron-positron ($\beta^+\beta^-$) pair production, and (3) at lower energies by photoabsorption. In these processes the energy of the photon is converted to the kinetic energy of the few electrons with which it interacts. Since electrons are much less penetrating than gamma-ray photons, their energy is largely trapped within the detector, where their ionizing effect creates a response convertible to an electrical output. In a gas-ionization device, such as Geiger counter, this occurs by the production of ion–electron pairs and in a solid-state device, such as a germanium detector, by production of electron-hole pairs. In a scintillation device, for example, a sodium iodide (NaI) detector, the response is caused by the emission of optical photons from atoms excited by the passage of energetic electrons. *See* COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION; GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER; SCINTILLATION COUNTER.

In accurate instruments the magnitude of the current pulse created by a single gamma-ray photon is closely proportional to the energy within the detector volume. However, gamma radiation is so penetrating that any particular event may involve only partial absorption of the photon. For example,

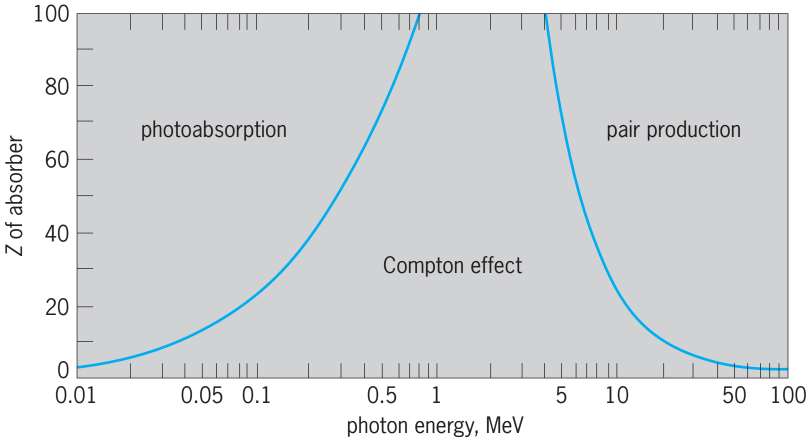


Fig. 1. Domains in photon energy versus atomic number (Z) of an absorber. Each of the three mechanisms by which gamma rays lose energy to ionization is of dominant importance in its domain. *(After R. D. Evans, The Atomic Nucleus, McGraw-Hill, 1955)*

a single Compton scattering may be followed by the escape of the scattered photon (now reduced in energy) from the detector, leaving behind only the energy of the scattered electron.

Photon interactions in matter. The three principal interactions of gamma rays with matter depend on the photon energy and the atomic number (Z) of the absorber. In photoabsorption, a photon is absorbed by collision with an atom, its energy and momentum being conserved by the ejection of an electron. This mechanism is important to gamma-ray detection since it is the only one of these interactions in which the photon energy is completely captured by ionization. This mechanism dominates at low photon energies and for absorbers of high Z; for high-energy photons, pair production dominates (**Fig. 1**). This process cannot occur unless the photon energy exceeds twice the rest-mass of an electron (2×511 keV), the minimum requirement for creating an electron-positron pair ($\beta^+ \beta^-$). Although this positron loses energy rapidly in a solid, it ultimately annihilates an electron to liberate two new gamma photons, each of energy 511 keV. To register the full energy, both photons must be contained within the detector.

Desirable detector characteristics are therefore (1) high Z, enhancing the domain of photoabsorption, (2) reasonable size, giving photons a chance to convert energy to ionization in more than one encounter, and (3) high density, maximizing the number of electrons encountered in the distance traveled.

Gamma-ray spectra and detector characteristics. An experiment may be imagined in which a detector is placed adjacent to a point source such as a cobalt-60 source that emits gamma rays of energies 1173 and 1332 keV. The output current pulses (one per photon detected) are amplified electronically and digitized. The resulting stream of data is stored in the form of a histogram showing numbers of events on the ordinate and output pulse amplitude on the abscissa, and called a spectrum by analogy with optical spectra. Since the source has only two photon energies (or wavelengths, photon energy E being related to wavelength λ by $\lambda = hc/E$, where h is Planck's constant and c is the speed of light), two lines are found in the spectrum (**Fig. 2**).

For events at the peak energy (or photopeak), all the photon energy is trapped within the sensitive volume. The background of partial-energy events is called the Compton platform. Such spectra help to define the three characteristic parameters of gamma-ray detectors: (1) resolution, the width of a peak at its half-height point, which measures the minimum energy separation at which two different photon peaks of comparable intensity are distinguishable; (2) total detection efficiency, the number of events registered in the spectrum versus the number of photons known to have struck the detector (by calculation from the source emission rate); and (3) peak-to-total ratio, the number of events in the photopeak versus the total number detected at that and all lower energies. These characteristics vary with photon energy, which must therefore be specified with the performance figures.

Gas-ionization detectors. These detectors are simple, cheap, and reliable but have very low efficiency because of the low density of their sensitive volume. Their principal application is in monitoring strong radiation fields for industry and radiation protection. The individual output pulses are electronically integrated; the instrument gives a reading of count rate or field intensity, often calibrated directly in dose

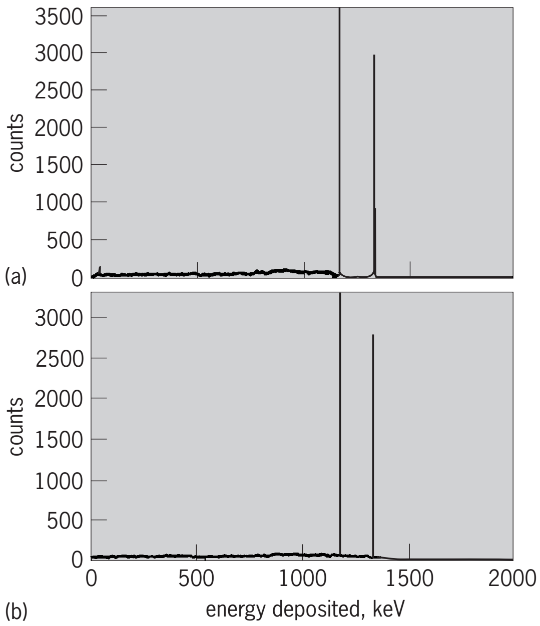


Fig. 2. Gamma-ray spectra of a cobalt-60 radiation source. (*a*) Spectrum from a high-purity germanium detector of 25% efficiency. The two peaks represent events in which the original photon energy (1173 and 1332 keV) was completely contained within the crystal. The platform at lower energies corresponds to events in which a photon lost some energy to the crystal but eventually escaped. (*b*) Spectrum that results from activating a bismuth germanate scintillation detector surrounding the high-purity germanium crystal so that the event is not recorded if both detectors give an output. The platform of degraded events seen in the high-purity germanium detector is substantially suppressed.

rate. The amplification occurs in the gas itself as a discharge develops under the intense electric field surrounding the anode wire. This process is virtually noise-free and results in a gain of about 10^6 on the original number of electrons liberated.

Scintillation detectors. These detectors employ certain materials which are nearly transparent and convert a significant amount of ionization energy into light emitted in the visible or near-visible region. The material is viewed by one or more photomultiplier tubes, which collect the light and convert it into an electrical pulse with very low noise and a typical gain of 10^8. The most common scintillator for gamma-ray application is sodium iodide with a thallium doping, NaI(Tl). Crystals can be produced in virtually any shape or size up to 15 × 15 in. (38 × 38 cm) cylinders; the most common laboratory detector is 3 × 3 in. (7.6 × 7.6 cm). For small crystals, the resolution is typically 6.5% at 662 keV (cesium-137 radiation) and increases approximately with the inverse square root of the energy. *See* PHOTOMULTIPLIER.

Other alkali halide scintillators in common laboratory use are barium fluoride (BaF_2) and cesium iodide (CsI). These have a higher density and average atomic number than sodium iodide, and for a given detector size give a higher efficiency and higher peak-to-total ratio, although their resolution is slightly poorer. The barium fluoride scintillator is used where a fast response is important, as with very high counting rates or fast electronic timing.

Another class of scintillator used for gamma-ray detection comprises the heavy-metal complex oxides, especially bismuth germanate ($Bi_4Ge_3O_{12}$). It has a very high density and offers the strongest gamma-radiation absorption of any known practical scintillator. In resolution it is appreciably poorer than sodium iodide; a good performance for a 3 × 3 in. (7.6 × 7.6 cm) crystal is 11.5% at 662 keV.

Solid-state detectors. For general application, the germanium detector is the most useful. The first successful ones, called Ge(Li) detectors, depended on lithium compensation of crystal imperfections. Manufacturing techniques have been improved, and detectors based on high-purity intrinsic germanium (HPGe), are now readily available. Germanium detectors function at liquid-nitrogen temperature. High-purity germanium detectors have the following advantages over Ge(Li) detectors: they may be thermally cycled, and radiation damage, such as that caused by neutrons, can be annealed cheaply and easily. *See* JUNCTION DETECTOR.

Germanium detectors have better resolution than most other gamma-ray detectors, which scales differently with energy than the resolution of scintillation detectors: a typical performance is 1 keV resolution at 100 keV, 2 keV at 1300 keV, and 3 keV at 2000 keV. Historically, the photopeak efficiency of germanium detectors has been defined in terms of a 3 × 3 in. (7.6 × 7.6 cm) NaI(Tl) detector at the same distance from a radiation source, usually cobalt-60. On this standard, an efficiency of 25% is typical for high-purity germanium, but detectors of up to 70% are routinely available, and a 100% detector has been demonstrated.

Silicon detectors have less general applicability. They must also be cooled to liquid-nitrogen temperature, and have significantly better resolution than germanium detectors but are less efficient. These detectors are useful in the spectroscopy of very low-energy radiation (below 50 keV) where the efficiency is equal but the resolution (typically 0.6 keV) nearly two times better than that of germanium.

Detector systems. Powerful instrumentation has been created by assembling separate gamma-ray detectors into a single unit. For example, Fig. 2*b* shows the performance gained by suppressing escaped radiation from a high-purity germanium detector. In this case the active crystal was surrounded by a bismuth germanate scintillator detector except in the direction of the radiation source. Electronic circuits processed the output pulses from both detectors so that if they were simultaneous (within 5×10^{-8} s) the event was rejected. In some instruments a spherical shell of scintillation detectors (typically 100 elements) surrounds the radiation source. In one application the outputs of all the detectors are summed whenever an event occurs. In this way the instrument acts as a gamma-ray calorimeter, registering an output proportional to the energy of gamma radiation emitted in the event. For example, the fusion of two heavy nuclei is typically accompanied by a burst of up to 30 gamma rays with a total energy of perhaps 30 MeV. Photon energies exceeding 100 MeV are associated with the decay of subnuclear systems such as the J/ψ particle. *See* ELEMENTARY PARTICLE; NUCLEAR FUSION; PARTICLE DETECTOR.

David Ward

Bibliography. D. A. Bromley (ed.), Detectors in nuclear science, *Nucl. Instrum. Meth.*, vol. 162, 1979; J. Cerny (ed.), *Nuclear Spectroscopy and Reactions*, pt. A, 1974; K. Debertin and R. G. Helmer, *Gamma and X-Ray Spectrometry with Semiconductor Detectors*, 1988; H. Ejiri and M. J. De Voigt, *Gamma-Ray and Electron Spectroscopy in Nuclear Physics*, 1989; R. M. Singro, *Introduction to Experimental Nuclear Physics,* 1975.

Gamma rays

Electromagnetic radiation emitted from excited atomic nuclei as an integral part of the process whereby the nucleus rearranges itself into a state of lower excitation (that is, energy content). For the theory of gamma emission *See* NUCLEAR STRUCTURE; RADIOACTIVITY.

Nature of gamma rays. The gamma ray is an electromagnetic radiation pulse—a photon—of very short wavelength. The electric (**E**) and magnetic (**H**) fields associated with the individual radiations oscillate in planes mutually perpendicular to each other and also the direction of propagation with a frequency ν which characterizes the energy of the radiation. The **E** and **H** fields exhibit various specified phase-and-amplitude relations, which define the character of the radiation as either electric (EL) or magnetic (ML). The second term in the designation indicates the order of the radiation as 2^L-pole, where the orders are monopole (2^0), dipole (2^1), quadrupole (2^2), and so on. The most common radiations are dipole and quadrupole. Gamma rays range in energy from a few kiloelectronvolts to 100 MeV, although most radiations are in the range 50–6000 keV. As such, they lie at the very upper high- frequency end of the family of electromagnetic radiations, which include also radio waves, light rays, and x-rays. *See* ELECTROMAGNETIC RADIATION; MULTIPOLE RADIATION; PHOTON.

Wave-particle duality. The dual nature of gamma rays is well understood in terms of the wavelike and particlelike behavior of the radiations. For a gamma ray of intrinsic frequency ν, the wavelength is $\lambda = c/\nu$, where c is the velocity of light; energy is $E = h\nu$, where h is Planck's constant. The photon has no rest mass or electric charge but, following the concept of mass-energy equivalence set forth by Einstein, has associated with it a momentum given by $p = h\nu/c = E/c$. *See* LIGHT; QUANTUM MECHANICS; RELATIVITY.

While gamma rays and x-rays are usually labeled by their energies (in kiloelectronvolts or megaelectronvolts), an equivalent useful specification in terms of either wavelength or frequency is easily obtained from the relationships given above. For less energetic members of the electromagnetic family, it is customary to classify radio waves in terms of wavelength (in meters) or frequency (in hertz). Photons in the visible and near-visible regime are labeled alternately by energy (in electronvolts) or wavelength (in nanometers). However, the general considerations on propagation, polarization, and Doppler shifts are quite similar for all. One of the early techniques for precise measurements of low-energy nuclear gamma rays utilized a bent-crystal spectrometer, which measures wavelength; the corresponding energies were subsequently obtained from the equivalence relations given above. *See* DOPPLER EFFECT; NUCLEAR CHEMISTRY; POLARIZATION OF WAVES.

Origin. From a historical perspective, the term "gamma ray" was introduced to label the electromagnetic radiations emanating from nuclear deexcitations (as distinct from alpha, beta, and x-rays). Early observations of nuclear giant-dipole resonance phenomena dealt with gamma rays in the range up to about 30 MeV. High-energy radiations are also observed in cosmic events; for example, the pi-zero meson (π^0) annihilates by the simultaneous emission of two gamma rays with total energies in the 100-MeV range. *See* GAMMA-RAY ASTRONOMY; GIANT NUCLEAR RESONANCES.

Other sources of very high energy electromagnetic radiations (or gamma rays) have become of increasing interest. Electron beams in the gigaelectronvolt range can produce gamma rays in the range of hundreds of megaelectronvolts via brehmsstrahlung from a fixed target, or alternatively from a collision between high-energy beam electrons and an opposing photon beam from a simple laser (all carried out in vacuum). *See* BREMSSTRAHLUNG; LASER; PARTICLE ACCELERATOR.

For example, if an argon laser [wavelength (λ) $\approx$ 350 nm; energy (E) $\approx$ 3.5 eV] is focused opposite to the electron beam direction, low-energy photons backscattered from the electrons (in Compton scattering) will achieve an energy boost proportional to the electron momentum, resulting in scattered photons with energies of around 500 MeV. With this technique, simultaneous observation of the scattered electron (direction and momentum) also defines the precise gamma-ray energy. Since the Compton scattering process conserves polarization, the polarization of the gamma-ray flux is also known. This production process has been used to map out the region of the nuclear delta resonance (at around 300 MeV), which lies well above the giant dipole resonance. *See* COMPTON EFFECT; DELTA RESONANCE.

Nuclear gamma rays. One of the most frequently utilized sources of nuclear gamma rays is ^{60}Co (that is, the cobalt isotope of $N = 33$ neutrons, $Z = 27$ protons, and thus of atomic mass number $A = N + Z = 60$). The decay process (**Fig. 1**) begins when ^{60}Co (in its ground state, or state of lowest possible excitation) decays to ^{60}Ni ($N = 32$, $Z = 28$) by the emission of a β^- particle. More than 99% of these decays lead to the 2506-keV level of ^{60}Ni; this level subsequently deexcites by an 1173-keV gamma transition to the 1332-keV level, which in turn emits a 1332-keV gamma ray leading to the ^{60}Ni ground state.

The gamma rays from ^{60}Ni carry information not only on the relative excitation of the ^{60}Ni levels, but also on the quantum-mechanical nature of the individual levels involved in the gamma decay. From the standpoint of nuclear physics, the levels of a given nucleus can be described most simply in terms of their excitation energies ($\mathbf{E}_x$) relative to the ground state, and in terms of the total angular momentum (**J**) and parity (π) quantum numbers given as $\mathbf{J}^\pi$. For a gamma-ray transition from initial state i to final state f, one obtains $E^i{}_x - E^f{}_x = E'_\gamma$ is the measured gamma energy after small (second-order) corrections

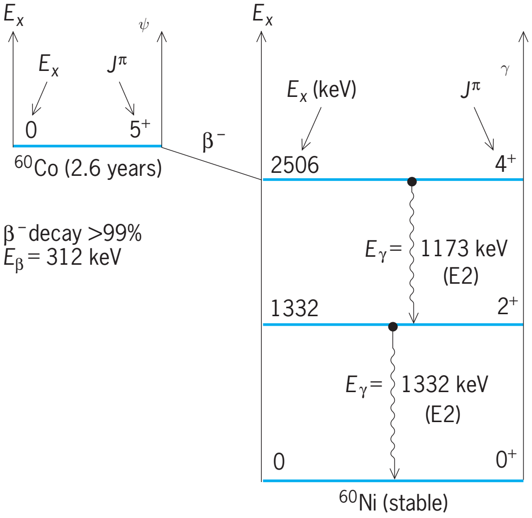


Fig. 1. Energy-level diagram illustrating the gamma decay of levels of ^{60}Ni resulting from beta decay of ^{60}Co.

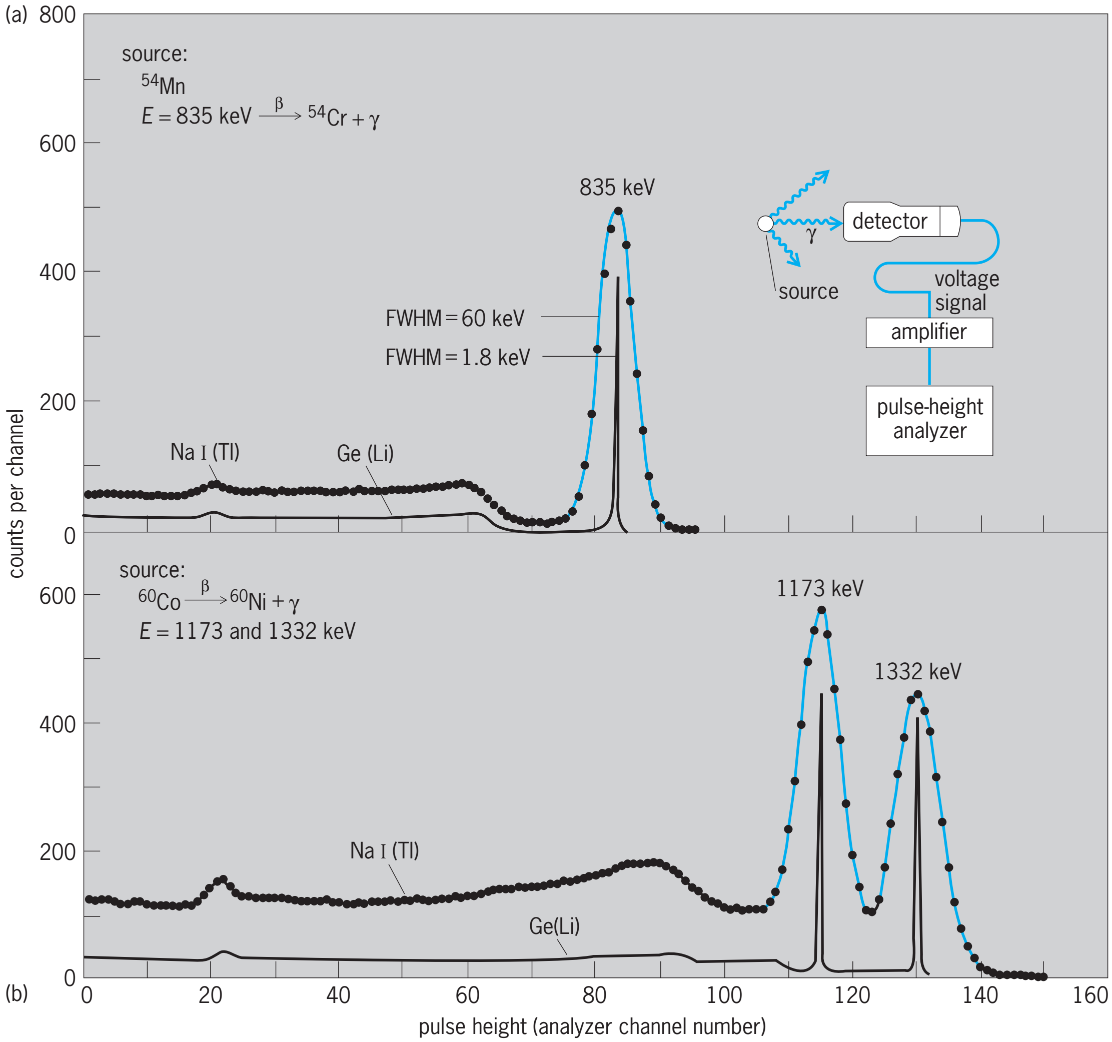


Fig. 2. Gamma-ray spectra from radioactive sources as measured with both NaI(Tl) and Ge(Li) detectors. Inset shows the components of detector apparatus. (*a*) ^{54}Mn source. (*b*) ^{60}Co sources.

for nuclear recoil and relativistic effects. Nuclear selection rules restrict the multipole character of the radiation according to the change in the quantum numbers J^π of the initial and final states. In Fig. 1, for example, the transitions must be electric quadrupole (E2), since they connect states of similar parity ($\pi = +$) by radiation of order $L \geq J_i - J_f = 2$. *See* NUCLEAR ISOMERISM; NUCLEAR SPECTRA; PARITY (QUANTUM MECHANICS); SELECTION RULES (PHYSICS); SPIN (QUANTUM MECHANICS).

Use as nuclear labels. Various nuclear species exhibit distinctly different nuclear configurations: the excited states, and thus the gamma rays which they produce, are also different. Precise measurements of the gamma-ray energies resulting from nuclear decays may therefore be used to identify the gamma-emitting nucleus, that is, not only the atomic number Z but also the specific isotope as designated by A. This has ramifications for nuclear research and also for a wide variety of more practical applications.

The two most widely used detectors for such studies are the NaI(Tl) detector and the Ge(Li) detector. **Figure 2** shows typical gamma spectra measured for sources of ^{60}Co and ^{54}Mn. Full-energy peaks are labeled by the gamma-ray energy, given in kiloelectronvolts. The figure of merit for these detectors, defined for a given gamma energy as the full-width-at-half-maximum (FWHM) for the full energy peak, is indicated. Although the more efficient NaI(Tl) detector can clearly distinguish the ^{60}Co and ^{54}Mn gamma rays, it is evident that the Ge(Li) detector, having a line width of only 1.8 keV, is more appropriate for complex nuclei, or for studies involving a greater number of source components. *See* GAMMA-RAY DETECTORS.

Applications to nuclear research. One of the most useful studies of the nucleus involves the bombardment of target nuclei by energetic nuclear projectiles, to form final nuclei in various excited states. For example, ^{48}Ca bombarded by ^{16}O makes ^{60}Ni strongly via the ^{48}Ca(^{16}O, $4n$)^{60}Ni reaction, as well as numerous other final species. Ge(Li) measurements of the decay gamma rays are routinely used to identify the various final nuclei according to their characteristic gamma rays, that is, the 1332- and 1173-keV gamma rays of ^{60}Ni, for example.

Precise measurements of the gamma energies, together with intensity and time-coincidence measurements, are then used to establish the sequence of gamma-ray decay, and thus construct from experimental evidence the nuclear level scheme. Angular correlation and linear polarization measurements determine the radiation character (as M1, E1, E2, or M2 or mixed) and thus the spin-parity of the nuclear levels. These studies provide a very useful tool for investigations of nuclear structure and classification of nuclear level schemes.

Practical applications. In these applications, the presence of gamma rays is used to detect the location or presence of radioactive atoms which have been deliberately introduced into the sample. In irradiation studies, for example, the sample is activated by placing it in the neutron flux from a reactor. The resultant gamma rays are identified according to isotope by Ge(Li) spectroscopy, and thus the composition of the original sample can be inferred. Such studies have been used to identify trace elements found as impurities in industrial production, or in ecological studies of the environment, such as minute quantities of tin or arsenic in plant and animal tissue. *See* ACTIVATION ANALYSIS.

In tracer studies, a small quantity of radioactive atoms is introduced into fluid systems (such as the human bloodstream), and the flow rate and diffusion can be mapped out by following the radioactivity. Local concentrations, as in tumors, can also be determined. *See* RADIOACTIVE TRACER.

Doppler shift. If $E_{\gamma 0} = h\nu_0$ is the gamma ray energy emitted by a nucleus at rest, then the energy $E_\gamma = h\nu$ emitted from a nucleus moving with velocity v at angle θ (with respect to the direction of motion) is given by Eq. (1) where c is the velocity of light. In

$$E_\gamma = E_{\gamma 0}\left(1 + \frac{v}{c}\cos\theta\right) \tag{1}$$

terms of the frequency ν, this expression is entirely analogous to the well-known Doppler shift of sound waves. Experimental measurements of the Doppler shift are used to determine the velocity of the nucleus and, more importantly, to shed light on the lifetime of the nuclear gamma-emitting state. A major advantage of this technique is that the same nuclear reaction which produces the excited nuclear states can also be employed to impart a large velocity to the nucleus.

For example, the velocity of ^{60}Ni nuclei produced via the ^{48}Ca(^{16}O,4n)^{60}Ni reaction at $E(^{16}\mathrm{O}) = 50$ MeV is $v/c = 0.00204$, and instead of $E_\gamma = 1332$ keV one should observe $E_\gamma = 1359$ keV. The extent of the shift is clearly within the resolving power of the Ge(Li) detector, which may therefore be used to measure E_γ and thus infer v. In most nuclear reactions, v is a known function of time t [that is, $v = v(t)$], and one therefore obtains a distribution of E_γ's whose precise shape may be related to the lifetime of the nuclear state.

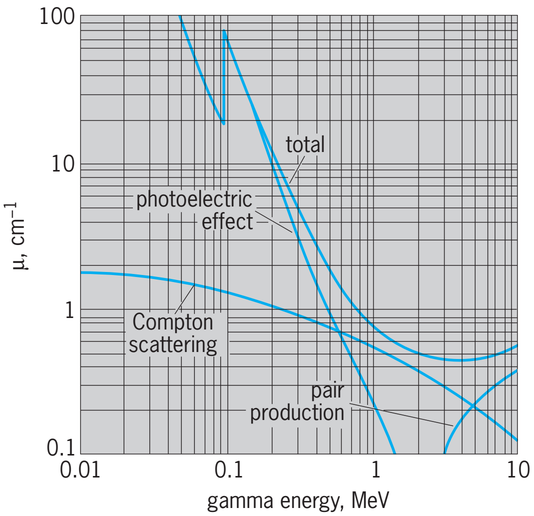


Fig. 3. **Graphic representation of partial and total attenuation coefficients for lead as a function of gamma energy. (*National Institute of Standards and Technology*)**

Doppler-shift measurements of gamma rays from recoil nuclei produced in nuclear reactions have been routinely used since the mid-1960s to measure nuclear lifetimes of 10^{-9} to 10^{-14} s, a range previously considered inaccessible to study.

Interaction with matter. For the three types of interaction with matter which together are responsible for the observable absorption of gamma rays, namely, Compton scattering, the photoelectric effect, and pair production, *See* COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION; PHOTOEMISSION.

The energy of a photon may be absorbed totally or partially in interaction with matter; in the latter case the energy of the photon is reduced and its direction of motion is changed. Photons are thus absorbed not gradually, but in discrete events, and one interaction is sufficient to remove a photon from a collimated beam of gamma rays. The intensity I of a beam decreases exponentially, as in Eq. (2), where x is the

$$I = I_0 e^{-\mu x} \tag{2}$$

path length, I_0 is the initial intensity, and μ is the linear attenuation coefficient, which is characteristic of the material and the gamma energy.

The dependence of the attenuation coefficient on gamma-ray energy is shown in **Fig. 3** for a lead absorber. For different absorbers, the attenuation is greater for the more dense materials. Most attenuation coefficients are tabulated as mass attenuation coefficients μ/ρ where ρ is the material or elemental density.

J. W. Olness

Bibliography. J. Cerny (ed.), *Nuclear Spectroscopy and Reactions*, pt. C, 1974; H. A. Enge and R. P. Redwine, *Introduction to Nuclear Physics*, 2d ed., 1995; R. D. Evans, *The Atomic Nucleus*, 1955, reprint 1982; K. S. Krane, *Introductory Nuclear Physics*, 1987.

Ganglion

A group of nerve cell bodies, usually located outside the brain and spinal cord. A ganglion located inside the central nervous system is called a nucleus.

The dorsal root ganglia are rounded clusters of cell bodies and fibers, surrounded by a connective tissue covering, located on the dorsal, or sensory, root of each spinal nerve. They lie just outside the cord and contain the cell bodies of sensory nerves whose fibers originate in some peripheral sensory receptor.

Other ganglia are given specific names which indicate their function or location, such as acoustic, cardiac, carotid, jugular, celiac, and sympathetic ganglia. Sympathetic ganglia, lying on either side of the vertebral column, unite by fiber strands to form a sympathetic chain. Preganglionic axons of the sympathetic system run along this path until they reach their terminal point in these ganglia, where they synapse with the postganglionic neurons. *See* SYMPATHETIC NERVOUS SYSTEM.

The term ganglion may be applied to a tumorlike, often cystic growth found on tendons, joints, and other connective tissues, but this usage is rare. *See* BRAIN; SPINAL CORD.

Walter Bock

Gangrene

A form of tissue death, or necrosis, usually occurring in an extremity and due to insufficient blood supply.

Dry gangrene. If no bacterial contamination is present, the part becomes dry, greenish-yellow, and finally turns brown or black. This is known as mummification. A sharp inflammatory border marks the edge of the adjacent viable tissue. This dry gangrene is seen most often in small portions of the extremities, such as the fingers and toes. Senile gangrene is the form caused by deterioration of blood supply in the elderly, usually as the result of progressive arteriosclerosis. Similar types are often present in diabetes mellitus, Reynaud's disease, and Buerger's disease (thromboangiitis obliterans).When tissue loss has progressed beyond the point of salvage, amputation may be required. *See* ARTERIOSCLEROSIS; DIABETES.

Moist or wet gangrene. When bacterial infection intervenes, putrefaction ensues, thus producing the moist or wet type of gangrene. Moist gangrene may occur in the body anywhere the blood supply is blocked and bacterial contamination occurs. The size of the area varies from a small, asymptomatic, local lesion to the destruction of an entire organ, such as the gall bladder or appendix. The parts most commonly affected by dry gangrene may also become contaminated, with the subsequent production of wet gangrene.

Gas gangrene. If bacterial contamination is present and the invading microorganisms are gas-producing *Clostridium* species, rapidly progressive gas gangrene is the result. This localized but rapidly spreading necrotizing wound infection is characterized by extensive edema with gas production and discoloration of the tissue, and is often contaminated by a putrefactive odor. The disease commonly arises following septic abortion (an abortion complicated by acute infection of the endometrium) or dirt contamination of deep wounds. The microbial flora of the gangrenous wound usually comprises one or more species of toxigenic anaerobic bacteria mixed with nontoxigenic anaerobic species, aerobic species, or both. The following gram-positive, spore-forming species are generally considered the primary causal agents of gas gangrene: *C. perfringens* (*C. welchii*), *C. septicum* (*Vibrion septique*), *C. novyi* (*C. oedematiens*), *C. histolyticum, C. fallax*, and *C. bifermentans*. Anaerobic streptococci may also cause gas gangrene.

Edward G. Stuart; N. Karle Mottet; Wilbert S. Aronow

Spread. Gangrene does not necessarily follow the presence of these organisms in a wound, as initiation of the disease depends on the virulence of the organisms and other factors relating to the resistance of the host. The virulence of the gangrene-producing organisms depends on the toxins produced, and the same species may produce several immunologically different toxins. Some of the toxins have been identified as specific enzymes, such as hyaluronidase, collagenase, and lecithinase. *See* TOXIN; VIRULENCE.

Diagnosis and treatment. The rapidly spreading nature of the disease precludes extensive laboratory diagnostic aids because therapeutic measures, possibly including amputation, usually must be instituted before laboratory results are available. Procedures should begin, however, for isolation of all species present and determination of toxin type and drug sensitivity. Initial thorough surgical debridement (removal of foreign material and devitalized tissue) of all contaminated wounds is essential. Polyvalent antitoxins are available for prophylactic and therapeutic use, and penicillin and tetracycline antibiotics are valuable in some cases. Formol toxoids and alum-precipitated formol toxoids for active immunization against the most important species have been produced but are not routinely used. The effectiveness of antibiotic therapy seems to depend on the species involved in the infection and the elapsed time between injury and treatment. Clostridial sepsis should be treated with intravenous clindamycin and high-dose penicillin. Hyperbaric oxygen therapy, which involves the administration of oxygen under greater-than-atmospheric pressure, is beneficial for individuals who are diagnosed early. *See* ANTIGEN; BIOLOGICALS; ENZYME; HYPERBARIC OXYGEN CHAMBER; TOXIN; TOXIN-ANTITOXIN REACTION.

Leland S. McClung; Wilbert S. Aronow

Bibliography. W. S. Aronow, Management of peripheral arterial disease, *Cardiol. Rev.*, 13:61–68, 2005; A. L. Bisno and D. L. Stevens, Streptococcal infections in skin and soft issues, *NEJM*, 334:240–245, 1996; N. Jallali, S. Withey and P. E. Butler, Hyperbaric oxygen as adjuvant therapy in the management of necrotizing fasciitis, *Amer. J. Surg.*, 189:462–466, 2005;

B. Lorber, Gas gangrene and other *Clostridium*-associated diseases, in G. L. Mandell et al. (eds.), *Principles and Practice of Infectious Diseases*, 5th ed., pp. 2549-2560, Churchill Livingstone, New York, 2000; B. N. Perry and W. E. Floyd III, Gas gangrene and necrotizing fasciitis in the upper extremity, *J. Surg. Orthop. Adv.*, 13:57–68, 2004; R. W. Titball, Gas gangrene: An open and closed case, *Microbiology*, 151(pt. 9):2821–2828, 2005.

Gantt chart

A graphic device that depicts tasks, machines, personnel, or whatever resources are required to accomplish a job on a calendar-oriented grid. Charts may be provided for various managerial levels and responsibilities, but detailed planning occurs at the lowest organizational level. Performance may be monitored and controlled throughout the organization.

Developed by Henry L. Gantt in 1917, the Gantt chart was the first scheduling model ever devised. The original intent of the chart was to assign and control jobs on factory machines, thereby minimizing delays and conflicts in scheduling. Machine loading could thus be planned weeks or months in advance.

Bar chart. An initial step in development of a Gantt chart may be to specify the tasks or activities making up a project, as shown in the **illustration**. The amount of time required for each activity is represented as a horizontal bar on the chart, with open triangles designating original start and finish dates in this example. The open start triangle is changed to a filled triangle upon inauguration of the activity, and the bar is filled in with vertical lines to indicate progress and completion. The open finish triangle is also filled upon completion. Slippage times are documented on the chart by broken lines, and the diamond symbols are employed to indicate rescheduled work. The vertical line on the chart is the current-date indicator and indicates present and future status of the project as of that date.

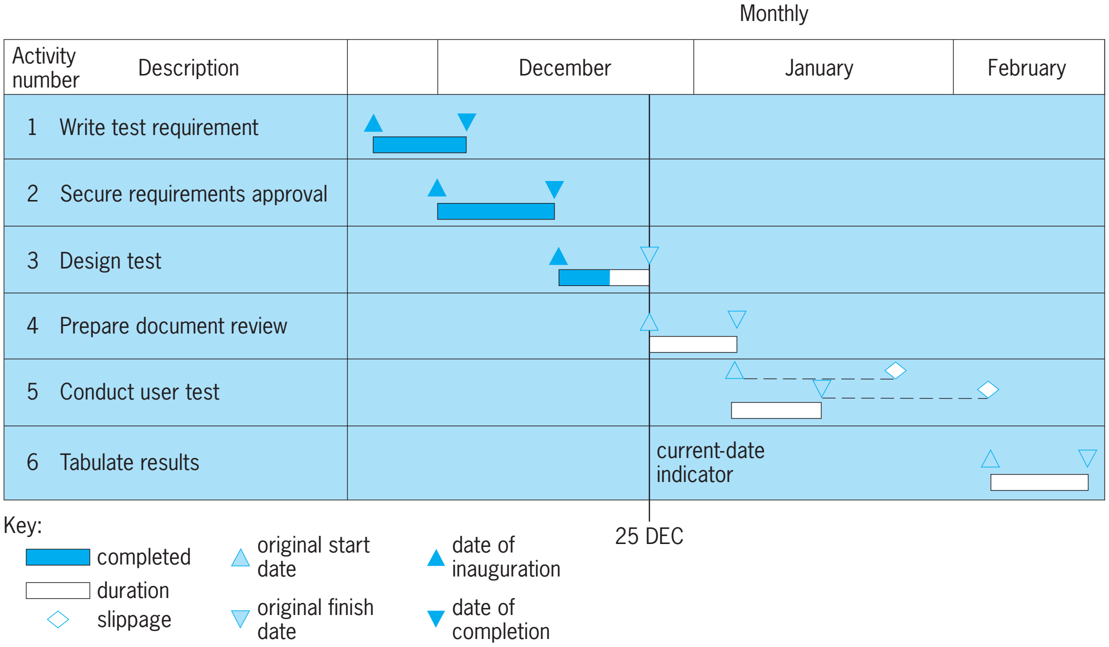
Updating of a Gantt chart will reveal difficulties encountered in the conduct of a project. Possible solutions include rescheduling, overtime, multishift operations, use of additional equipment and facilities, and changes in method.

Color coding may be used to isolate or identify specified tasks. Charts may designate individual workers, machines, and facilities rather than the tasks shown in the illustration. The time scale may be subdivided into days, weeks, or other units.

Computer programs have become available for drawing and updating Gantt charts. Large and elaborate commercial display boards using a pegboard design are available, some employing revolving disks, tape, or plastic slides to portray the status of planned and current work.

The Gantt chart is an effective tool for planning and scheduling operations involving a minimum of dependencies and interrelationships among the activities. The technique is best applied to activities for which time durations are not difficult to estimate, since there is no provision for treatment of uncertainty. On the other hand, the charts are easy to construct and understand, even though they may contain a great amount of information. In general, the charts are easily maintained provided the task requirements are somewhat static.

Milestone chart. An outgrowth of the bar chart technique is the milestone chart. A milestone is an important activity in the sequence of project completion. The most significant activities may be designated major milestones. The primary difference in utilization of this concept is the graphic display, since the method and collection of data are the same.



Example of a Gantt bar chart.

The milestone approach offers no intrinsic improvement over the basic Gantt chart but provides a means for focusing resources on critical items. *See* PERT.

Lawrence S. Hill

Bibliography. Harvard Business Review Staff, *Project Management*, 1991; E. Menipaz, *Essentials of Production and Operations Management*, 1984; G. Salvendy, *Handbook of Industrial Engineering*, 2d ed., 1991.

Garlic

A hardy perennial, *Allium sativum*, of Asiatic origin and belonging to the plant order Liliales. Garlic is grown for its pungent bulbs, segments of which are used primarily for seasoning. Europeans have grown it for more than 200 years. Propagation is commonly by bulb segments, sometimes called cloves; seeds are seldom produced. Cultural practices are similar to those used for onions. Popular varieties are Italian, Tahiti, and Creole or Mexican. Harvest of the mature dry bulbs is 7–8 months after planting. Garlic salt is made from dehydrated cloves. California is the most important producing state; smaller acreages are planted to garlic in Louisiana and Texas. *See* LILIALES; ONION.

J. John Carew

Garnet

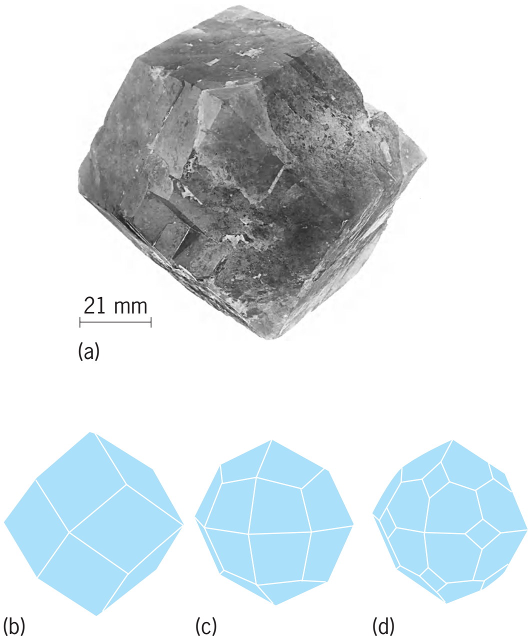
A hard, dense silicate mineral which occurs as crystals of cubic symmetry in a wide range of geologic environments. The general chemical formula of the silicate garnet group is $A_3B_2(SiO_4)_3$ where, in natural occurrences, the A cations are dominantly Fe^{2}, Mn^{2}, Mg, and Ca and the B cations are Al, Fe^{3}, and Cr^{3}.

The garnet crystal structure is based on isolated (SiO_4) tetrahedra with the A cations surrounded by eight oxygens at the corners of a cube and the B cations in octahedral oxygen coordination. Crystals are usually well developed with the most common forms being the dodecahedron and the trapezohedron (see **illustration**).

Chemical compositions. The garnet mineral group is generally divided into a number of individual species on the basis of chemical composition. The names of the more common of these species and the idealized compositions to which they refer are as follows:

A series	*B series*
Pyrope $Mg_3Al_2(SiO_4)_3$	Grossular $Ca_3Al_2(SiO_4)_3$
Almandine $Fe_3Al_2(SiO_4)_3$	Andradite $Ca_3Fe_2(SiO_4)_3$
Spessartin $Mn_3Al_2(SiO_4)_3$	Uvarovite $Ca_3Cr_2(SiO_4)_3$

Wide ranges of solid solution exist between the end members of series A and within series B, with specific garnets generally being named after the dominant species they contain. Solid solution between the two series is, however, restricted in nature partly due to the instability of such garnets and partly because of the scarcity of rocks with compositions appropriate to their formation. All compositions between pyrope and grossular and between almandine and grossular may be synthesized stably at elevated pressures (20–40 kilobars or 24 gigapascals). All of the pure end members with the exception of pyrope may be synthesized at low pressure; pressures on the order of 12 kb (1.2 GPa) are required to stabilize pyrope.



Garnet. (*a*) Grossular, Lake Jaco, Mexico (*American Museum of Natural History specimens*). (*b–d*) Crystal habits (*after C. Klein and C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*). (*b*) Dodecahedron. (*c*) Trapezohedron. (*d*) Combined dodecahedron and trapezohedron.

Extensive substitutions involving ions other than those listed above are not common, but there are occurrences of $(OH)_4$ replacing (SiO_4) tetrahedra in grossular (hydrogrossular) and of Ti substitution in andradite. In the latter case there is complete solid solution between andradite and $Ca_3Ti_2Fe_2SiO_{12}$ (schorlomite). *See* SOLID SOLUTION.

Physical properties. Garnets are substantially denser than most chemically analogous silicates, with specific gravities ranging between 3.58 (pyrope) and 4.32 (almandine). They also have high refractive indices (1.71–1.89) and hardness, on Mohs scale, of $6^{1}/_{2}$ to $7^{1}/_{2}$. The relative hardness, coupled with the absence of cleavage, has led to the use of garnet as an abrasive.

The color of garnet is primarily controlled by its chemical composition. Common garnets of series A are red, reddish brown, or black, while grossular-andradite mixtures are white or yellowish brown to black depending on iron content. Uvarovite is emerald green. Gem varieties of garnet are generally clear, deep red pyrope, but some darker almandines and andradites are also used as gemstones. Demantoid is a rare grass green to emerald green gem variety of andradite. Garnets are normally optically isotropic, although some grossular-andradites exhibit biaxial optical anomalies of a sectoral nature. *See* GEM; SILICATE MINERALS.

The cubic space group of garnet is Ia3d with unit cell edge of between 1.146 and 1.205 nanometers depending on composition. Cell dimensions vary linearly with composition within the two main series. The complete range of stable solid solubility between Mg and Ca end members and between Fe^2 and Ca end members is unique among the major silicate mineral groups, limited miscibility being the general rule. The region of negative deviation from linearity at low calcium concentration, coupled with thermodynamic data, indicates that the substitution of Ca for Mg and Fe^2 may be stabilized by the formation of an ordered Ca(Mg,Fe) compound of lower symmetry than Ia3d.

Occurrence. Garnets are widespread in their occurrence, particularly in rocks which formed at high temperatures and pressures. Among the A series, pure pyrope is stable only at depths greater than about 24 mi (40 km), and garnets rich in this component are thus confined to rocks of subcrustal origin. Pyropic garnets occur in ultrabasic rocks, particularly those which occur as xenoliths in kimberlite, in serpentinized peridotites, and in eclogites. Almandine and spessartine are stable at much lower pressures and hence occur in crustal rocks. In particular, Fe^2-rich garnets are present in many gneissic and schistose metamorphic rocks formed by the recrystallization of argillaceous sediments at depths greater than about 6 mi (10 km). Such garnets typically contain about 75% almandine together with small amounts of spessartine, pyrope, and grossular. During progressive metamorphism of argillaceous rocks, garnet appears at the expense of chlorite when temperatures reach about 930°F (500°C). Because of the large, readily identifiable crystals which form, the first appearance of garnet (at the garnet "isograd") is commonly used by geologists as an index of the intensity, or grade, of metamorphism. Spessartine-rich garnets occur in metamorphosed sedimentary manganese deposits and in some granite pegmatites and acid igneous rocks. In the latter case the garnets may also be rich in almandine component.

Grossularite and andradite are typical minerals of the low-pressure metamorphism of impure calcareous rocks. Solid solutions between these end members are also common constituents of metasomatic skarns, where they are typically associated with magnetite and iron-bearing pyroxenes. Uvarovite-rich garnets are rare, occurring chiefly in serpentinites, where they are typically associated with chromite.

Garnets are strongly resistant to weathering and alteration and are hence widespread constituents of sands and sediments in areas of garnetiferous primary rocks. *See* METAMORPHIC ROCKS; SILICATE MINERALS.

B. J. Wood

Bibliography. L. G. Berry and B. Mason, *Mineralogy*, 2d ed., 1983; C. Klein and C. S. Hurlbut, *Dana's Manual of Mineralogy*, 21st ed., 1993, revised 1998; R. C. Newton, T. V. Charlu, and O. J. Kleppa, Thermochemistry of high pressure garnets and clinopyroxenes in the system CaO-MgO-Al_2O_3-SiO_2, *Geochim. Cosmochim. Acta*, 41:369–377, 1977.

Gas

A phase of matter characterized by relatively low density, high fluidity, and lack of rigidity. A gas expands readily to fill any containing vessel. Usually a small change of pressure or temperature produces a large change in the volume of the gas. The equation of state describes the relation between the pressure, volume, and temperature of the gas. In contrast to a crystal, the molecules in a gas have no long-range order.

At sufficiently high temperatures and sufficiently low pressures, all substances obey the ideal gas, or perfect gas, equation of state, shown as Eq. (1),

$$p\overline{V} = RT \qquad (1)$$

where p is the pressure, T is the absolute temperature, $\overline{V}$ is the molar volume, and R is the gas constant. Absolute temperature T expressed on the Kelvin scale is related to temperature t expressed on the Celsius scale as in Eq. (2).

$$T = t + 273.16 \qquad (2)$$

The gas constant is

$$\begin{aligned} R &= 82.0567 \text{ cm}^3\text{-atm/(mole)(K)} \\ &= 82.0544 \text{ ml-atm/(mole)(K)} \end{aligned}$$

The molar volume is the molecular weight divided by the gas density.

Empirical equations of state. At lower temperatures and higher pressures, the equation of state of a real gas deviates from that of a perfect gas. Various empirical relations have been proposed to explain the behavior of real gases. The equations of J. van der Waals (1899), Eq. (3), of P. E. M. Berthelot (1907), Eq. (4), and F. Dieterici (1899), Eq. (5), are frequently used.

$$\left(p + \frac{a}{\overline{V}^2}\right)(\overline{V} - b) = RT \qquad (3)$$

$$\left(p + \frac{a}{T\overline{V}_2}\right)(\overline{V} - b) = RT \qquad (4)$$

$$pe^{a/\overline{V}RT}(\overline{V} - b) = RT \qquad (5)$$

In these equations, a and b are constants characteristic of the particular substance under considerations. In a qualitative sense, b is the excluded volume due to the finite size of the molecules and roughly equal to four times the volume of 1 mole of molecules. The constant a represents the effect of the forces of attraction between the molecules. In particular, the internal energy of a van der Waals gas is $-a\overline{V}_i$. None of these relations gives a good representation of the compressibility of real gases over a wide range of temperature and pressure. However, they reproduce qualitatively the leading features of experimental pressure-volume-temperature surfaces.

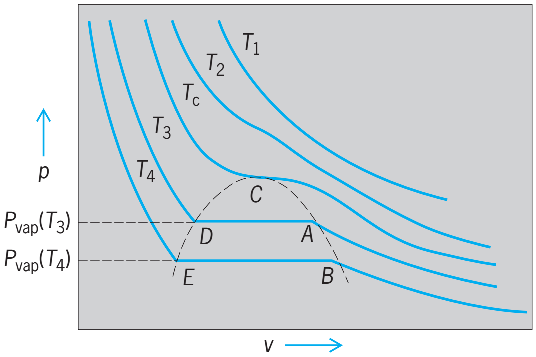


Fig. 1. Schematic isotherms of a real gas. *C* is the critical point. Points *A* and *B* give the volume of gas in equilibrium with the liquid phase at their respective vapor pressures. Similarly, *D* and *E* are the volumes of liquid in equilibrium with the gas phase.

Schematic isotherms of a real gas, or curves showing the pressure as a function of the volume for fixed values of the temperature, are shown in **Fig. 1**. Here T_1 is a very high temperature and its isotherm deviates only slightly from that of a perfect gas; T_2 is a somewhat lower temperature where the deviations from the perfect gas equation are quite large; and T_c is the critical temperature. The critical temperature is the highest temperature at which a liquid can exist. That is, at temperatures equal to or greater than the critical temperature, the gas phase is the only phase that can exist (at equilibrium) regardless of the pressure. Along the isotherm for T_c lies the critical point, *C*, which is characterized by zero first and second partial derivatives of the pressure with respect to the volume. This is expressed as Eq. (6).

$$\left(\frac{\partial p}{\partial \overline{V}}\right)_c = \left(\frac{\partial^2 p}{\partial \overline{V}}\right)_c = 0 \qquad (6)$$

At temperatures lower than the critical, such as T_3 or T_4, the equilibrium isotherms have a discontinuous slope at the vapor pressure. At pressures less than the vapor pressure, the substance is gaseous; at pressures greater than the vapor pressure, the substance is liquid; at the vapor pressure, the gas and liquid phases (separated by an interface) exist in equilibrium.

Along one of the isotherms of the empirical equations of state discussed above, the first and second derivatives of the pressure with respect to the volume are zero. The location of this critical point in terms of the constants *a* and *b* is shown below; p_c and $\overline{V}_c$ are the pressure and volume at the critical temperature.

	van der Waals	*Berthelot*	*Dieterici*
p_c	$\frac{a}{27b^2}$	$\left(\frac{aR}{216b^3}\right)^{1/2}$	$\frac{a}{4e^2b^2}$
$\overline{V}_c$	$3b$	$3b$	$2b$
T_c	$\frac{8a}{27Rb}$	$\left(\frac{8a}{27Rb}\right)^{1/2}$	$\frac{a}{4Rb}$
$\frac{p_c\overline{V}_c}{RT_c}$	0.3750	0.3750	0.2706

Some typical values of $p_c\overline{V}_c/RT_c$ for real gases are as follows: 0.30 for the noble gases, 0.27 for most of the hydrocarbons, 0.243 for ammonia, and 0.232 for water. The van der Waals and Berthelot equations of state, Eqs. (3) and (4), cannot quantitatively reproduce the critical behavior of real gases because no substance has a value of $p_c\overline{V}_c/RT_c$ as large as 0.375. The Dieterici equation, Eq. (5), gives a good representation of the critical region for the light hydrocarbons but does not represent well the noble gases or water.

At temperatures lower than the critical point, the analytical equations of state, such as the van der Waals, Berthelot, or Dieterici equations, give S-shaped isotherms as shown in **Fig. 2**. From thermodynamic considerations, the vapor pressure is determined by the requirement that the cross-hatched area *DEO* be equal to the cross-hatched area *AOB*. Under equilibrium conditions, the portion of this isotherm lying between *A* and *D* cannot occur. However, if a gas is suddenly compressed, points along the segment *AB* may be realized for a short period until enough condensation nuclei form to create the liquid phase. Similarly, if a liquid is suddenly overexpanded, points along *DE* may occur for a short time. For low temperatures, the point *E* may represent a negative pressure corresponding to the tensile strength of the liquid. However, the simple analytical equations of state cannot be used for a quantitative estimate of these transient phenomena. Actually, it is easy to show that the van der Waals, Berthelot, and Dieterici equations give poor representations of the liquid phase since the volume of most liquids (near their freezing point) is considerably less than the constant *b*.

Principle of corresponding states. In the early studies, it was observed that the equations of state of many substances are qualitatively similar and can be correlated by a simple scaling of the variables. To describe this result, the reduced or dimensionless

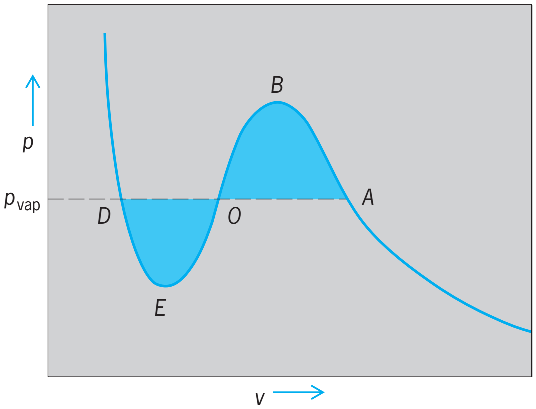


Fig. 2. Schematic low-temperature isotherm as given by van der Waals, Berthelot, or Dieterici equations of state. Here the line *DOA* corresponds to the vapor pressure. The point *A* gives the volume of the gas in equilibrium with the liquid phase, and *D* gives the volume of the liquid. The segment of the curve *DE* represents overexpansion of the liquid. The segment *AB* corresponds to supersaturation of the vapor. However, the segment *EOB* could not be attained experimentally.

variables, indicated by a subscript r, are defined by dividing each variable by its value at the critical point. These variables are given in Eqs. (7)–(9).

$$p_r = \frac{p}{p_c} \tag{7}$$

$$T_r = \frac{T}{T_c} \tag{8}$$

$$\overline{V}_r = \frac{\overline{V}}{\overline{V}_c} \tag{9}$$

In its most elementary form, the principle of corresponding states asserts that the reduced pressure, p_r, is the same function of the reduced volume and temperature, v_r and T_r, for all substances. An immediate consequence of this statement is the statement that the compressibility factor, expressed as Eq. (10),

$$z = \frac{p\overline{V}}{RT} \tag{10}$$

is a universal function of the reduced pressure and the reduced temperature. This principle is the basis of the generalized compressibility chart shown in **Fig. 3**. This chart was derived from data on the equation-of-state behavior of a number of common gases.

It follows directly from the principle of corresponding states that the compressibility factor at the critical point z_c should be a universal constant. It is found experimentally that this constant varies somewhat from one substance to another. On this account, empirical tables were developed of the compressibility factor and other thermodynamic properties of gases as functions of the reduced pressure and reduced temperature for a range of values of z_c. Such generalized corresponding-states treatments are very useful in predicting the behavior of a substance on the basis of scant experimental data.

Theoretical considerations. The equation-of-state behavior of a substance is closely related to the manner in which the constituent molecules interact. Through statistical mechanical considerations, it is possible to obtain some information about this relationship. If the molecules are spherically symmetrical, the force acting between a pair of molecules depends only on r, the distance between them. It is then convenient to describe this interaction by means of the intermolecular potential $\varphi(r)$ defined so that the force is the negative of the derivative of $\varphi(r)$ with respect to r. *See* INTERMOLECULAR FORCES.

Two theoretical approaches to the equation of state have been developed. In one of these approaches, the pressure is expressed in terms of the partition function Z and the total volume V of the container in the manner of Eq. (11). Here k is

$$p = kT\frac{\partial \ln Z}{\partial V} \tag{11}$$

the Boltzmann constant or the gas constant divided by the Avogadro number N_0, $k = R/N_0$. For a gas made up of spherical molecules or atoms with no internal structure, the partition function is given as Eq. (12) In this expression, φ_{ij} is the energy of in-

$$Z = \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{3N/2} \times \int \exp\left(-\sum_{i>j}\frac{\varphi_{ij}}{kT}\right) d\overline{V}_1 d\overline{V}_2 \cdots d\overline{V}_N \tag{12}$$

teraction of molecules i and j and the summation is over all pairs of molecules, h is Planck's constant, N is the total number of molecules, and the integration is over the three cartesian coordinates of each of the N molecules. The expression for the partition function may easily be generalized to include the effects of the structure of the molecules and the effects of quantum mechanics.

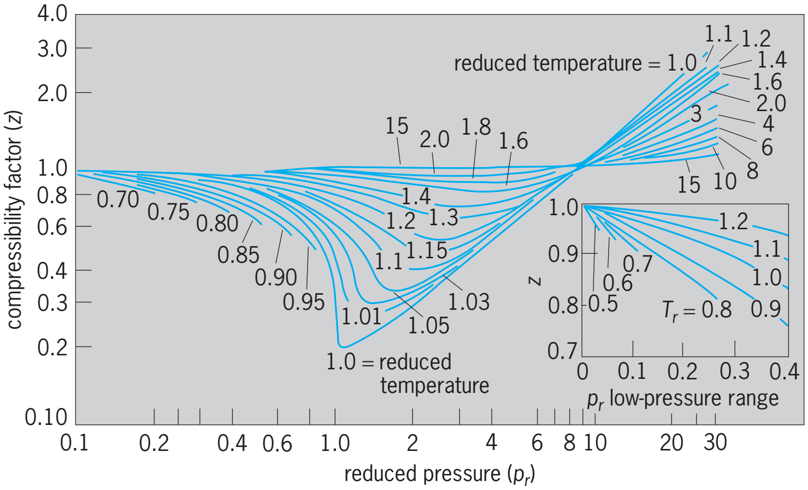


Fig. 3. The compressibility factor $p\overline{V}_i/RT$ as a function of the reduced pressure $p_r = p/p_c$ and reduced temperature $T_r = T/T_c$. (*After O. A. Hougen, K. M. Watson, and R. A. Ragatz, Chemical Process Principles, pt. 2, Wiley, 1959*)

In another theoretical approach to the equation of state, the pressure may be written as Eq. (13), where

$$p = \frac{NkT}{V} - \frac{2\pi N^2}{3V^2}\int g(r)\frac{d\varphi}{dr}r^3\,dr \tag{13}$$

$g(r)$ is the radial distribution function. This function is defined by the statement that $2\pi(N_2/V)g(r)r_2dr$ is the number of pairs of molecules in the gas for which the separation distance lies between r and $r + dr$. The radial distribution function may be determined experimentally by the scattering of x-rays. Theoretical expressions for $g(r)$ are being developed.

The compressibility factor $z = pV/NkT$ may be considered as a function of the temperature, T, and the molar volume, $\overline{V}$. In the virial form of the equation of state, z is expressed as a series expansion in inverse powers of $\overline{V}$, as in Eq. (14). The coefficients

$$z = 1 + \frac{B(T)}{\overline{V}} + \frac{C(T)}{\overline{V}^3} + \cdots \tag{14}$$

$B(T)$, $C(T)$, . . . , which are functions of the temperature, are referred to as the second, third, . . . , virial coefficients. This expansion is an important method of representing the deviations from ideal gas behavior. From statistical mechanics, the virial coefficients can be expressed in terms of the intermolecular potential. In particular, the second virial coefficient is Eq. (15). If the intermolecular potential is known,

$$B(T) = 2\pi N_0 \int (1 - e^{-\varphi/kT})\, r^2\, dr \qquad (15)$$

Eq. (15) provides a convenient method of predicting the first-order deviation of the gas from perfect gas behavior. The relation has often been used in the reverse manner to obtain information about the intermolecular potential. Often $\varphi(r)$ is expressed in the Lennard-Jones (6-12) form, Eq. (16), where ϵ and σ

$$\varphi(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] \qquad (16)$$

are constants characteristic of a particular substance. Values of these constants for many substances have been tabulated. In terms of these constants, the second virial coefficient has the form of Eq. (17), where

$$B(T) = \frac{2}{3} \pi N_0 \sigma^3 B^* \frac{kT}{\epsilon} \qquad (17)$$

$B^*(kT/\epsilon)$ is a universal function. If all substances obeyed this Lennard-Jones (6-12) potential, the simple form of the law of corresponding states would be rigorously correct. *See* CHEMICAL THERMODYNAMICS; KINETIC THEORY OF MATTER.

C. F. Curtiss; J. O. Hirschfelder

Bibliography. I. B. Cohen (ed.), *Laws of Gases*, 1981; J. O. Hirschfelder et al., *Molecular Theory of Gases and Liquids*, 1964; R. Holub and P. Vonka, *The Chemical Equilibrium of Gaseous Systems*, 1976; I. N. Levine, *Physical Chemistry*, 4th ed., 1995; R. Mohilla and B. Ferencz, *Chemical Process Dynamics*, 1982.

Gas absorption operations

The separation of solute gases from gaseous mixtures of noncondensables by transfer into a liquid solvent. This recovery is achieved by contacting the gas stream with a liquid that offers specific or selective solubility for the solute gas or gases to be recovered. The operation of absorption is applied in industry to purify process streams or recover valuable components of the stream. It is used extensively to remove toxic or noxious components (pollutants) from effluent gas streams.

Examples of application of gas absorption are the removal of hydrogen sulfide and mercaptans from natural gas, the recovery of carbon monoxide in petrochemical syntheses, the removal of sulfur oxides from power plant stack gases, and the recovery of silicon tetrafluoride and hydrogen fluoride from fertilizer production stack gases.

The absorption process requires the following steps: (1) diffusion of the solute gas molecules through the host gas to the liquid boundary layer based on a concentration gradient, (2) solvation of the solute gas in the host liquid based on gas-liquid solubility, and (3) diffusion of the solute gas based on concentration gradient, thus depleting the liquid boundary layer and permitting further solvation. The removal of the solute gas from the boundary layer is often accomplished by adding neutralizing agents to the host liquid to change the molecular form of the solute gas. This process is called absorption accompanied by chemical reaction.

Factors affecting the kinetics of the gas absorption process

Zone	Factor
Gas	Gas velocity
	Molecular weights and sizes of host and solute gases
	Temperature
	Concentration gradient
Solvation	Solubility in host liquid
Liquid	Molecular weights and sizes of host liquid and solute gas
	Temperature
	Viscosity of liquid
	Liquid surface velocity
	Surface renewal rate
	Concentration gradient of solute gas
	Concentration gradient of neutralizing reagent

The factors affecting the rate of absorption or the kinetics of transfer from the host gas to the host liquid are indicated in the **table**.

Historically, the kinetics of gas absorpion have been represented by statistical average overall coefficients in the form of Eq. (1), where N_A is the rate of

$$N_A = K_{G\alpha} V (P - P^*) \qquad (1)$$

mass-transfer quantity of solute gas transferred per

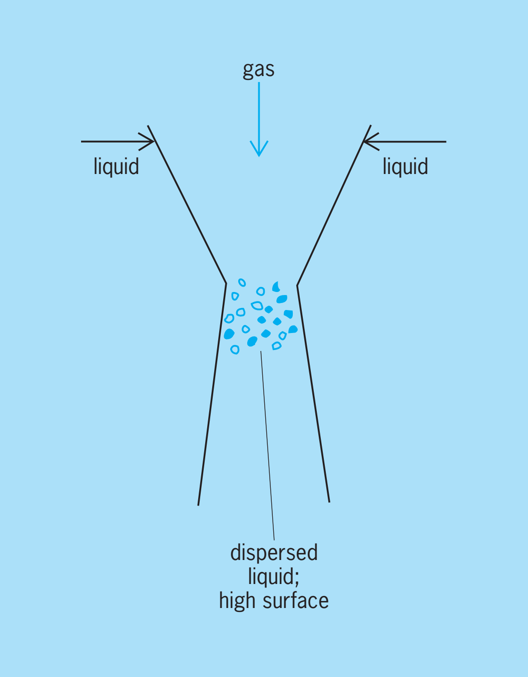


Fig. 1. Operation of a venturi system.

unit time; $K_{G\alpha}$ is the mass-transfer coefficient based on partial pressure gradient from the host-liquid bulk concentrations, quantity per unit time per volume of the transfer system per unit partial pressure gradient; V is volume of the transfer system; P is partial pressure of the solute gas in the bulk of the host gas; and P^* is partial pressure of the host gas in equilibrium with the bulk of the host liquid.

An alternative method of sizing the system is the transfer unit concept expressed in Eqs. (2)–(4),

$$N_{oG} = \int \frac{dp}{P - f(P^*)} \frac{1 - f(P_f)}{1 - P} \tag{2}$$

$$H_{oG} = \frac{Gm}{K_{G\alpha\pi}} \frac{1 - f(P_f)}{1 - P} \tag{3}$$

$$H = N_{oG} H_{oG} \tag{4}$$

where H is height of the absorber; N_{oG} is number of transfer units; H_{oG} is height of a transfer unit; $K_{G\alpha}$, P, and P^* are as in Eq. (1) ; P_f is partial pressure of the solute gas in the interfacial film between gas and liquid; π is total pressure of operation; and Gm is molar gas flow rate.

The effect of the relative resistances in the liquid and gas phases are established by the relationship shown in Eq. (5), where $K_{G\alpha}$ is as in Eq. (1); $k_{G\alpha}$ is

$$K_{G\alpha} = \frac{1}{\dfrac{1}{k_{G\alpha}} + \dfrac{m}{k_{L\alpha}}} \tag{5}$$

mass-transfer coefficient in the gas phase only, quantity per unit per volume of transfer system per unit partial pressure gradient; $k_{L\alpha}$ is mass-transfer coefficient in liquid phase only, quantity per unit time per unit volume per unit concentration gradient; and m is partial pressure–liquid phase concentration relationship of solute gas.

The gas-phase and liquid-phase mass-transfer coefficients are established as Eqs. (6) and (7), where D_G

$$k_G = \frac{D_G \pi}{RT \chi P_{BM}} \tag{6}$$

$$k_L = Z \left(\frac{D_L \phi}{\pi t} \right)^{1/2} \tag{7}$$

and D_L are diffusivities in the gas and liquid phases, respectively; R is the universal gas constant; T is the absolute temperature; χ is the effective thickness of the gas boundary layer; P_{BM} is the partial pressure of the inert gases; ϕ is the surface renewal rate; and t is time of contact.

It has been established that except for unique situations of very high solubility, the transfer in the liquid phase tends to become the controlling factor in rate of absorption.

The emphasis in design of systems has been to increase the surface renewal rate and to increase the effective liquid diffusivity by neutralizing the solute gas in the liquid phase as rapidly as possible. In the case of neutralization or total removal, the mass-transfer coefficient in the liquid phase follows the relationship shown in Eq. (8), where k'_L is liquid-phase mass-

$$\frac{k'_L}{k_L} = 1 + \left(\frac{q}{C_{ai}} \frac{D_B}{D_L} \right)^n \tag{8}$$

transfer coefficient when chemical reaction occurs; q is concentration of unreacted reagent in the bulk of the host liquid; C_{ai} is concentration of solute gas in the liquid phase (physically absorbed) in equilibrium with P; D_B and D_L are diffusivities of the solute

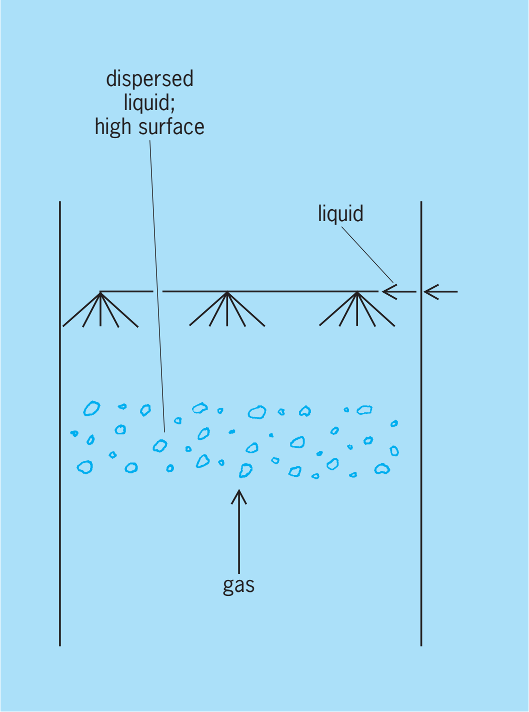


Fig. 2. Operation of a spray tower.

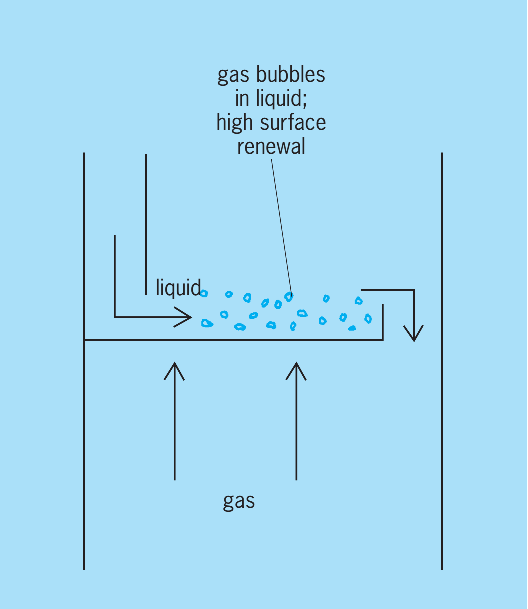


Fig. 3. Operation of a tray tower.

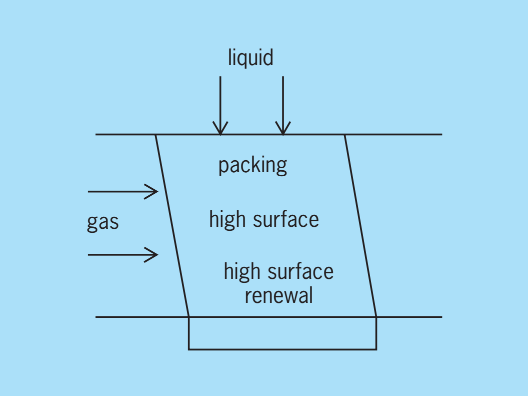


Fig. 4. Operation of a packed tower.

gas and reagent in the host liquid, respectively; and *n* is an exponent, generally less than 1. The transfer systems require effective contact of liquid and gas, with maximum surface contact and surface renewal, and minimum consumption of energy. The types of equipment used for this service are venturis (high surface, no surface renewal); spray towers (high surface, no surface renewal); tray towers (high surface renewal); and packed towers (high surface, high surface renewal).

The venturi system (**Fig. 1**), requiring high energy input, is often used where large quantities of particulates are present in the gas or the scrubbing liquid has large concentrations of suspended solids. The spray tower (**Fig. 2**) is used where the transfer requirements are low. The tray tower (**Fig. 3**) is not used extensively because of poor efficiency in low-concentration gas systems. The packed tower (**Fig. 4**) is the predominant system used in gas absorption because of its inherent differential mechanism, low power consumption, and capability for handling low concentrations of particulates in gas streams. It is often used in conjunction with a solids recovery unit such as a venturi.

The packed column volume generally is randomly filled with a commercially available geometric shape, ranging from 1 to 4 in. (2.5 to 10 cm) in the largest dimension, forms that provide either large surfaces or capability for surface renewal, or both.

The gas absorption equipment, usually called scrubbers, generally use water or solutions of reagents in the water as the host liquid, although hydrocarbon liquids are used for specific applications in the chemical industry. The efficiency of absorption can be uniquely high, especially in the case of absorption accompanied by chemical reaction. Reduction of the solute gas concentration to approximately 1–5 ppm is often required for air-pollution control. These effluent concentrations have been achieved in air-pollution control in the steel, fertilizer, glass, and pulp and paper industries. The gas absorption equipment range in capacity from 100 cubic feet per minute (4.7×10^{-2} m^3/s) to 2,000,000 cfm (940 m^3/s). *See* ADSORPTION; DISTILLATION; HUMIDIFICATION.

Aaron J. Teller

Gas and atmosphere analysis

Qualitative identifications and quantitative determinations of substances essential for the evaluation of the air quality in the ambient air and in the industrial workplace.

The terms air pollution and gas analysis are used here in the broad sense. Air pollution refers to unwanted air contaminants in both the ambient air and the work environment, since many of the same pollutants are found at both places and are assessed using similar procedures. These pollutants may exist in the gaseous state or as aerosols. Aerosols are liquid droplets or solid particles small enough (0.01 to 100 micrometers) to remain dispersed in the air for extensive periods of time. Gas analysis refers to the analysis of both gases and vapors. The term vapor is used for the gaseous phase of substances that are liquids or solids at standard temperature and pressure. Thus the gaseous states of gasoline, mercury, and water are examples of vapors, whereas methane, hydrogen, and ozone are gases. Many important pollutants are vapors that have arisen from the volatilization or decomposition of organic materials. *See* AEROSOL; AIR POLLUTION; INDUSTRIAL HEALTH AND SAFETY.

Qualitative Identification

The qualitative identification of air pollutants can be extremely complex, and may require the use of several instruments which provide complementary information about composition and structure. Since the entire sample is often limited to milligram or microgram quantities, the classical identification methods, such as boiling point and refractive index determinations, functional group tests, combustion analyses, and derivative preparations, have been largely replaced by instrumental methods. Information for identification purposes is now generally obtained from instruments such as mass, nuclear magnetic resonance, infrared, and ultraviolet spectrometers that rely upon the response of a molecule to an energy probe.

Organic compounds. These are identified by means of mass spectroscopy and gas chromatography.

Mass spectroscopy. This is probably the single most powerful technique for the qualitative identification of volatile organic compounds, and has been particularly useful in the identification of many environmental contaminants. When a sample is introduced into the mass spectrometer, electron bombardment causes the parent molecule to lose an electron and form a positive ion. Some of the parent ions are also fragmented into characteristic daughter ions, while other ions remain intact. All of the ions are accelerated, separated, and focused on an ion detector by means of either a magnetic field or a quadrupole mass analyzer. Using microgram quantities of pure materials, the mass spectrometer yields information about the molecular weight and the presence of other atoms, such as nitrogen, oxygen, and halogens, within the molecule. In addition, the fragmentation pattern often provides a unique so-called fingerprint of a molecule, allowing positive identification. If the gas is a mixture, interpretation of the mass spectral data is difficult since the fragmentation patterns are superimposed. However, interfacing the mass spectrometer to a gas chromatograph provides an elegant solution to this problem. *See* MASS SPECTROMETRY.

Gas chromatography. A gas chromatograph is essentially a highly efficient apparatus for separating a complex mixture into individual components. When a mixture of components is injected into a gas chromatograph equipped with an appropriate column, the components travel down the column at different rates and therefore reach the end of the column at different times. The mass spectrometer located at the end of the column can then analyze each component separately as it leaves the column. In essence, the gas chromatograph allows the mass spectrometer to analyze a complex mixture as a series of pure components. More than 100 compounds have been identified and quantified in automobile exhaust by using a gas chromatograph–mass spectrometer combination. *See* GAS CHROMATOGRAPHY.

Metals. A powerful technique for identifying metals is inductively coupled plasma spectrometry. An inductively coupled plasma spectrometer is an emission spectrometer capable of operating at 6000–8000 K (10,000–14,000°F). Emission spectrometers raise the temperature of the sample to a sufficiently high temperature that chemical bonds are broken and the electrons are raised from their stable energy levels to higher energy levels. The excited atoms lose a part of the excess energy by emitting a characteristic wavelength of light. Sensors placed around the plasma can identify and quantify 35 metals in a mixture. *See* EMISSION SPECTROCHEMICAL ANALYSIS.

Quantitative Analysis

Once a qualitative identification of an important pollutant has been established, further interest often centers on quantifying the levels of the pollutant as a function of time at various sites.

Methods. The methods employed chiefly for quantification can be classified for convenience into direct and indirect procedures. Direct-reading instruments are generally portable and may analyze and display their results in a few seconds or minutes, and can operate in a continuous or semicontinuous mode. Indirect methods are those involving collection and storage of a sample for subsequent analysis. Both direct and indirect methods have inherent advantages and disadvantages. By using indirect methods, samples with several pollutants can be simultaneously collected from a number of different sites with relatively inexpensive collection devices and later analyzed at a central laboratory. On the other hand, direct methods may require one instrument for each pollutant at each sampling site, and thus may become prohibitively expensive. However, reduction of the delay before the results are available may be the basis for selecting a direct over an indirect method, or the pollutant in question may not be stable under the conditions of storage. For example, if a worker needs to enter an enclosure with an atmosphere potentially hazardous to life because of an oxygen deficiency, the presence of an explosive mixture, or the presence of a high concentration of a toxic gas, it is essential to have the analytical results available in a few minutes.

Direct methods. These consist of methods utilizing colorimetric indicating devices and instrumental methods.

Colorimetric indicators. Three types of direct-reading colorimetric indicators have been utilized: liquid reagents, chemically treated papers, and glass tubes containing solid chemicals (detector tubes). The simplest of these methods is the detector tube. Detector tubes are constructed by filling a glass tube with silica gel coated with color-forming chemicals. For use, the ends of the sealed tube are broken and a specific volume of air, typically 6 in.^{3} (100 cm^3), is drawn through the tube at a controlled rate. Detector tubes often utilize the same color-forming chemical to detect several different gases, and therefore may be nonspecific for mixtures of these gases. Temperature, humidity, age, and uniformity of packing also influence the performance. Detector tubes for analyzing approximately 400 different gases are commercially available. Accuracy is sometimes low, and detector tubes for only 25 gases meet the National Institute for Occupational Safety and Health (NIOSH) accuracy requirement of ±25%. For some gases, semicontinuous analyzers have been developed that operate by pulling a fixed volume of air through a paper tape impregnated with a color-forming reagent. The intensity of the color is then measured for quantification. Phosgene, arsine, hydrogen sulfide, nitric oxide, chlorine, and toluene diisocyanate have been analyzed by indicating tapes. *See* COLORIMETRY.

Direct-reading instruments. With the availability of stable and sensitive electronics, direct-reading instruments capable of measuring gases directly at the parts-per-billion range were developed. Most direct-reading instruments contain a sampling system, electronics for processing signals, a portable power supply, a display system, and a detector. The detector or sensor is a component that is capable of converting some characteristic property of the gas into an electrical signal. While there are dozens of properties for the bases of operation of these detectors, the most sensitive and popular detectors are based on electrical or thermal conductivity, ultraviolet or infrared absorption, mass spectrometry, electron capture, flame ionization, flame photometry, heat of combustion, and chemiluminescence. Many of these detectors respond to the presence of 10^{-9} g quantities, and even to 10^{-12} g levels. In addition to improved accuracy, precision, and analysis time, another advantage is that most instruments produce an electrical signal which can be fed into a computer for process control, averaging, and record keeping. Rapid fluctuations and hourly, daily, and yearly averages are readily obtained. Several important instruments are based on the following principles.

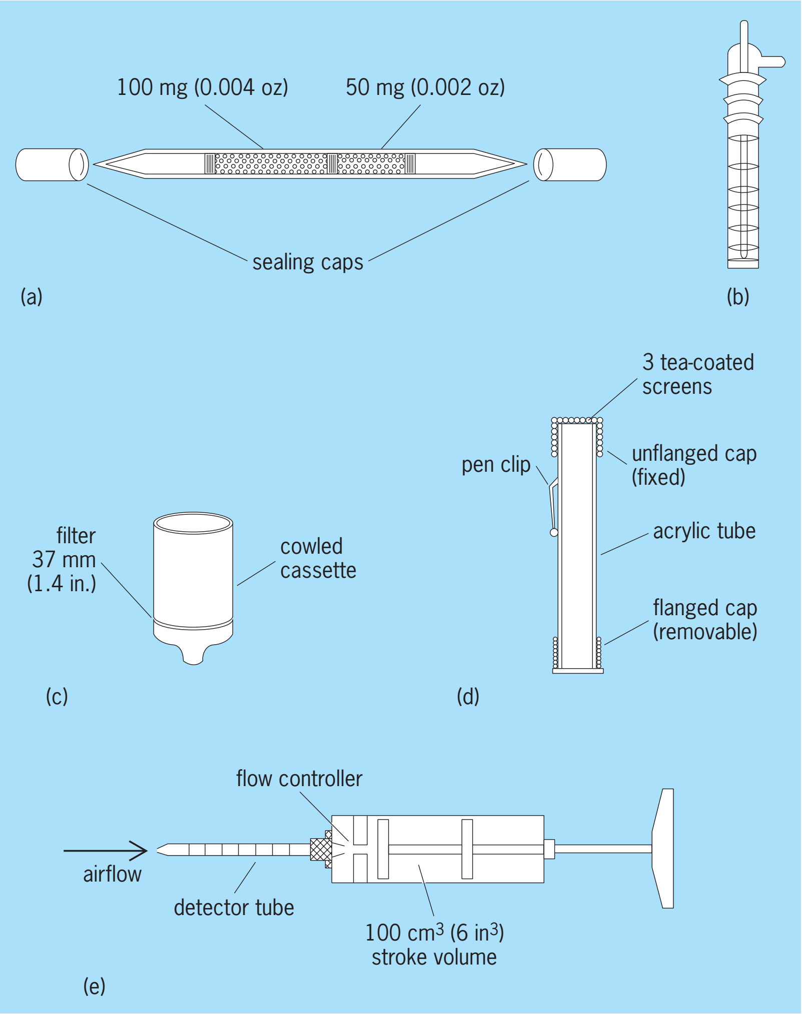


Fig. 1. Collection devices for monitoring air in the workplace. (*a*) Solid sorbent (charcoal) sampling tube. (*b*) Midget impinger. (c) Cowled cassette and its filter. (*d*) Passive monitor, used for nitrogen dioxide. (*e*) Piston-type detector tube pump.

1. *Heat of combustion*. Many portable direct-reading meters for explosive atmospheres are based on the principle of catalytic or controlled combustion on a heated filament. The filament is usually one arm of a Wheatstone bridge circuit. The resulting heat of combustion changes the resistance of the filament (usually platinum), and the resulting imbalance is related to the concentration of the gas. A meter displays the results. This method is nonspecific, and gives an indication of all combustible gases present in the range from about 100 parts per million (ppm) to a few percent.

2. *Chemiluminescence*. The phenomenon of chemiluminescence is employed for the determination of levels of ozone, oxides of nitrogen, and sulfur compounds. Chemiluminescence is an emissive process that occurs when all or part of the energy of a chemical reaction is released as light rather than heat. A familiar example is the “cold light” of fireflies. Ozone levels in the range from 1 to 100 ppb can be determined by measuring the emission at 585 nanometers, which occurs when ozone is mixed with excess ethylene. Similarly, nitric oxide (NO) levels from 10 ppb to 5000 ppm can be measured by a chemiluminescence method. The analysis of nitrogen dioxide (NO_2) is performed by reducing NO_2 to NO in a catalytic converter and measuring the NO emission. Mixtures of NO and NO_2 can be analyzed by measuring the NO level by chemiluminescence, reducing the NO_2 to NO, and again measuring the chemiluminescence level. The NO_2 level is obtained by difference. Measurement of the chemiluminescence produced by thermal means in a hydrogen-rich flame allows the detection of sulfur compounds from 1 to 1000 parts per billion (ppb). With the use of appropriate light filters, interferences by other gases can be reduced.

3. *Infrared absorption*. Measurement of carbon monoxide utilizes a variety of other principles, including infrared absorption at a fixed wavelength, nondispersive infrared absorption, measurement of the heat of reaction in oxidizing carbon monoxide to carbon dioxide, and gas chromatography.

4. *Gas chromatography*. The gas chromatograph is widely used both for the analysis of collected samples and as a semicontinuous direct-reading instrument. The availability of many different detectors capable of measuring the effluent from the gas-chromatograph column and the development of valves for automatic injection of samples and for directing sample flow have further extended the versatility of this instrument. Accessories are available that can collect samples from about 12 different distant sites through sampling lines for sequential injection directly into a gas chromatograph for analysis. Another application involves the use of valves to inject an air sample into a gas chromatograph equipped with a flame ionization detector to measure the methane content. A second air sample is injected directly into the detector to measure the total hydrocarbon content. The methane signal is subtracted from the total hydrocarbon content to provide an indication of the smog potential.

5. *Light scattering*. Portable dust monitors using light-scattering principles can directly measure concentrations to 0.01 mg/m^3. The sampling period can be varied but is typically about 1 min. The results are displayed as the average value for the sampling period. Some instruments can distinguish between dust particles and fibers. These have had some success in measuring airborne asbestos fibers.

6. *Electrochemical sensors*. Gases such as oxygen, carbon monoxide, carbon dioxide, sulfur dioxide, and hydrogen sulfide can be measured by using electrochemical sensors. Usually a separate sensor is required for each gas, but several can be combined into a single instrument.

7. *Ionization detectors*. Flame ionization detectors and photoionization detectors are capable of measuring many organic compounds in the parts-per-million range. These detectors produce ions, and a current is produced in proportion to the type of molecule and the number of the molecules. Flame ionization detectors use a hydrogen flame to produce the ions, and the photoionization detectors use ultraviolet light to produce ions. *See* CHEMILUMINESCENCE; LIGHT-SCATTERING PHOTOMETRY.

Indirect methods. For indirect methods, the main collection devices are freeze traps, bubblers, evacuated bulbs, plastic bags, and solid sorbents. Because of their convenience, solid sorbents dominate collection procedures. NIOSH developed a versatile method for industrially important vapors, based on the sorption of the vapors on activated charcoal or, to a lesser extent, on other solid sorbents such as silica gel and porous polymers. Typically, in this technique a few liters of air are pulled through a glass tube containing about 0.004 oz (100 mg) of charcoal. The charcoal tube is only 7 cm × 6 mm (3 in. × 0.2 in.), and has the advantage that it can be placed on the worker's lapel. A battery-operated pump small enough to fit into a shirt pocket is connected by a plastic tube to the collecting device, so that the contaminants are continuously collected from the breathing zone of the worker. Many solvent vapors and gases are efficiently trapped and held on the charcoal. The ends of the sample tube are then capped, and the tube is returned to a laboratory for analysis. In the laboratory the tube is broken open, and the charcoal poured into carbon disulfide to desorb the trapped vapors. Following desorption, a sample of the solution is injected into a gas chromatograph for quantification.

This technique has been highly successful for several classes of compounds, such as aromatics, aliphatics, alcohols, esters, aldehydes, and chlorinated compounds. Sulfur- and nitrogen-containing compounds can also be analyzed by using a gas chromatograph which is equipped with a sulfur- or nitrogen-sensitive detector.

Applications. The major applications of gas and atmosphere analysis involve industrial hygiene, indoor air quality, hazardous waste sites and spills, and sampling of ambient air and stationary sources.

Industrial hygiene. The Occupational Safety and Health Administration (OSHA) has established exposure standards for about 600 substances. Because a number of factors affect exposure, sampling at the breathing zone of the worker is preferred. Filters, sorbent tubes, impingers, and passive monitors are used to collect the samples (**Fig. 1**). Both OSHA and NIOSH have developed a number of sampling and analytical methods (**Table 1**).

Indoor air quality. Complaints about the indoor air quality became prevalent in the 1970s. NIOSH

TABLE 1. Examples of NIOSH analytical methods

Chemical[*]	Flow rate, liters/min	Minimum volume, liters	Maximum volume, liters	Analytical technique	Sampler	Accuracy[†] (±), %	Validation status[‡]
Ketones (6, for example, acetone)	0.01–0.2	0.5	3	Gas chromatography–flame ionization detector	Coconut shell charcoal tube	6–8	Full
Inorganic acids (6, for example, hydrofluoric)	0.2–0.5	3	100	Ion chromatography	Silica gel	12–23	Full
Asbestos and other fibers	0.5–16	400	Variable	Phase contrast microscopy	Mixed cellulose ester fiber	46	Full (special criteria)
Metals (26, for example, cadmium)	1–4	Varies ∽5	Varies ∽200	Inductively coupled plasma–atomic emission spectroscopy	Mixed cellulose ester fiber	Varies with metal	Partial
Carbon dioxide	0.02–0.1	0.5	80% of bag capacity	Gas chromatography–thermal conductivity detector	Bag	5.3	Full
Cresol	0.01–0.1	1	24	Gas chromatography–flame ionization detector	Porous polymer	Not determined	Partial
Chlorine	0.3–1	2	90	Ion chromatography–conductivity detector	Silver membrane filter	14.8	Full
Ethylenimine	0.2	1	48	High-performance liquid chromatography–ultraviolet	Bubbler	15.5	Full
Formaldehyde	0.2–1	1	100	Visible absorption spectrophotometry	Impinger	18	Full
Nitroglycerin	0.2–1.0	3	100	Gas chromatography–electron capture detector	Porous polymer	20.3	Full
Nitrogen dioxide	0.025–0.2	1.5	6	Visible absorption spectrophotometry	Treated solid sorbent	14.6	Full
Organophosphorous pesticides	0.2–1	12	Varies ∽240	Gas chromatography–flame photometric detector	Treated solid sorbent for pesticides	Varies ∽18	Full
Particulates, nuisance dust	1–2	7	133	Gravimetric	Polyvinyl chloride filter	11.0	Full
Silica, crystalline, respirable	1.7	400	1000	X-ray diffraction	Cyclone and polyvinyl chloride filter	18	Full
Toluene	Diffusion controlled	(15 min)	(8 h)	Gas chromatography–flame ionization detector	Passive	7.4	Full

[*] Methods are shown for specific chemicals or families of chemicals. The number of chemicals for which the method is applicable is sometimes given.

[†] The overall accuracy, including the sampling and analytical error, is expressed as the percent at the 95% confidence level.

[‡] Validation may be expressed as Full, Partial, or Unrated. Full validation means that the method has been tested with generated samples and is within ±25% of the true value.

SOURCE: *After NIOSH Manual of Analytical Methods*, 4th ed., August 1994.

TABLE 2. EPA sampling and analytical methodologies

Compound	Sampling methodology	Analytical procedure
Total particulate matter	Filtration	Gravimetric
Respirable particulate matter	Cyclones with filtration	Gravimetric
	Cascade impaction with filtration	Gravimetric
Nitrogen oxides	Absorption	Spectrophotometric or ion chromatographic
Sulfur dioxide	Direct analysis absorption	Chemiluminscence
		Titration
	Direct analysis	Ultraviolet, nondispersive infrared or fluorescence
Carbon monoxide	Flexible or rigid container	Nondispersive infrared or spectrophotometric
Acid gases	Direct analysis absorption	Specific ion electrode, ion chromatography, or colorimetric
Heavy metals	Filtration with absorption	Atomic absorption spectrophotometric
Volatile organics	Flexible/rigid containers, adsorption, absorption, cryogenic sampling, direct analysis	GC/FID, GC/MS GC/PID, GC/ECD, GC/FPD, GC/MS*
Semivolatile organics	Filtration with adsorption	GC/FID, GC/MS GC/PID, GC/ECD, GC/FPD, GC/MS*

*CEM = cellulose ester membrane. GC = gas chromatography. ICAP = inductively coupled argon plasma. FID = flame ionization detector. PID = photoionization detector. MS = mass spectrometry. FPD = flame photometric detector. ECD = electron capture detector.

conducted over 500 investigations and reported the following causes: building materials contamination (4%), microbiological contamination (5%), outside contamination (10%), inside contamination (15%), ventilation inadequate (53%), and unknown (13%). Ventilation accounted for over 50% of the complaints. In addition to checking the ventilation system, NIOSH found that the concentration of carbon dioxide could be used as an indicator of the adequacy of the ventilation. Carbon dioxide is exhaled by humans; although the level of carbon dioxide is not the cause of the complaints, it serves as an indicator of the buildup of all air pollutants emitted by people, such as perfumes and shaving lotions. When the levels of carbon dioxide increase from the outside levels (250–300 ppm) to above 800 ppm, complaints about the indoor air quality begin. At 1200 ppm, many people will complain. *See* AIR POLLUTION.

Ambient air and stationary source sampling. In many sampling situations, regulatory requirements dictate the type of sampling and analytical methodologies required during ambient air and stationary source sampling programs (**Table 2**). However, even in those situations, variables such as sampling location, sampling time, sampling procedures, gas stream characteristics, chemical properties, analytical detection limits, analytical interferences, and test program accuracy and precision requirements must be evaluated to ensure that applicable quality assurance objectives are satisfied. Sampling techniques must provide for representative samples with minimal deterioration and contamination. In addition, the collected sample must be of sufficient quantity to ensure that the concentration is above the detection limit of the analytical method.

The procedures employed for ambient air and source sampling are essentially the same procedures as described for industrial hygiene. The main differences are that larger volumes of the air are sampled and higher flow rates are used. One other major difference that is not encountered in industrial hygiene sampling is isokinetic sampling, which is required for particulate matter in stacks.

Gas sampling may be accomplished with evacuated flasks (**Fig. 2**), bags, canisters, an impinger sampling system, and sorbent tube sampling techniques. This method is limited by the size of

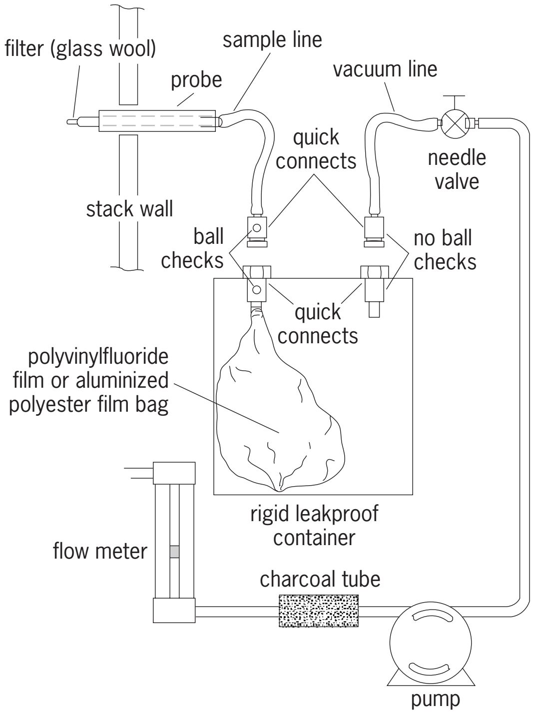


Fig. 2. Evacuated-flask sampling system.

sample collected and possible nonintegration of sample. William R. Burg; Michael Glowatz, Jr.; Laurence S. Bernson

Bibliography. American Conference of Governmental Industrial Hygienists, *Air Sampling Instruments for Evaluation of Atmospheric Contaminants*, 9th ed., 2001; M. J. Boss and D. W. Day, *Air Sampling and Industrial Hygiene Engineering*, 2000; National Institute for Occupational Safety and Health, *Manual of Analytical Methods*, vols. 1 and 2, DHHS Pub. 84–100, 1993; G. D. Wight, *Fundamentals of Air Sampling*, 1994; E. D. Winegar, *Sampling and Analysis of Airborne Pollutants*, 1993.

Gas chromatography

A method for the separation and analysis of complex mixtures of volatile organic and inorganic compounds. Most compounds with boiling points less than about 250°C (480°F) can be readily analyzed by this technique. A complex mixture is separated into its components by eluting the components from a heated column packed with sorbent by means of a moving-gas phase. A typical chromatogram is shown in **Fig. 1**.

Classification. Gas chromatography may be classified into two major divisions: gas-liquid chromatography, where the sorbent is a nonvolatile liquid called the stationary-liquid phase, coated as a thin layer on an inert, granular solid support, and gas-solid chromatography, where the sorbent is a granular solid of large surface area. The moving-gas phase, called the carrier gas, is an inert gas such as nitrogen or helium which flows through the chromatographic column packed with the sorbent. The solute partitions, or divides, itself between the moving-gas phase and the sorbent and moves through the column at a rate dependent upon its partition coefficient, or solubility, in the liquid phase (gas-liquid chromatography) or upon its adsorption coefficient on the packing (gas-solid chromatography) and the carrier-gas flow rate. Open tubular glass or stainless steel capillary tubes of 0.005–0.02 in. (0.1–0.5 mm) inside diameter and length often as great as 300 ft (90 m), coated on the inside walls with a nonvolatile stationary liquid phase, are also widely used in the separation of complex mixtures.

Gas-solid chromatography is historically the older of the two forms of gas chromatography but was formerly limited primarily to the separation of permanent gases or relatively nonpolar solutes of low molecular weight, whereas gas-liquid chromatography has been the more popular and more versatile approach to the separation of a wide range of higher-molecular-weight compounds, owing to the large choice of liquid phases available. Research in gas-solid absorption chromatography, however, has shown considerable progress in decreasing objectionable peak tailing, thus providing higher separation efficiencies than were formerly possible. It is entirely possible that gas-solid chromatography will achieve greater status in future years. *See* ADSORPTION.

Basic apparatus. The apparatus used in gas chromatography consists of four basic components: a carrier-gas supply and flow controller, a sample inlet system providing a means for introduction of the sample, the chromatographic column and associated column oven, and the detector system. A schematic diagram of a working instrument is shown in **Fig. 2**.

Carrier-gas and flow monitor. Although the carrier gas is most commonly nitrogen or helium, other gases such as carbon dioxide, argon, xenon, and hydrogen are occasionally used. If the detector is a thermal conductivity detector, then helium is the carrier gas of choice. This is because helium has the largest thermal conductivity of all gases with the exception of hydrogen, a flammable and explosive gas. Since the

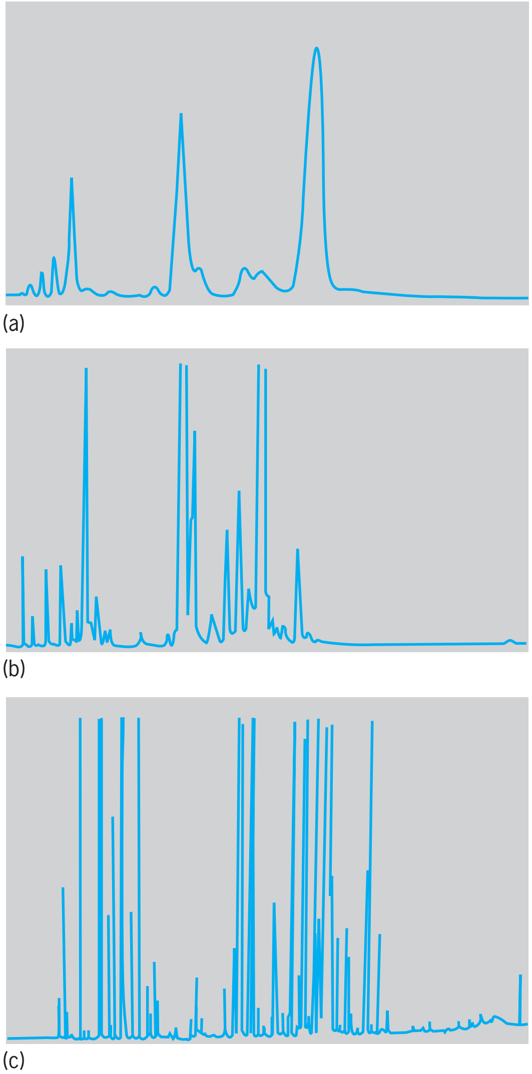


Fig. 1. Three generations in gas chromatography. Peppermint oil separated on (*a*) a 0.25 in. × 6 ft (6 mm × 1.8 m) packed column; (*b*) a 0.03 in. × 500 ft (0.8 mm × 150 m) stainless steel capillary column; and (*c*) a 0.98 in. × 164 ft (0.25 mm × 50 m) glass capillary column. (*After W. Jennings, The uses of glass capillary columns for food and essential oil analysis, J. Chromatog. Sci., 17:636–639, 1979*)

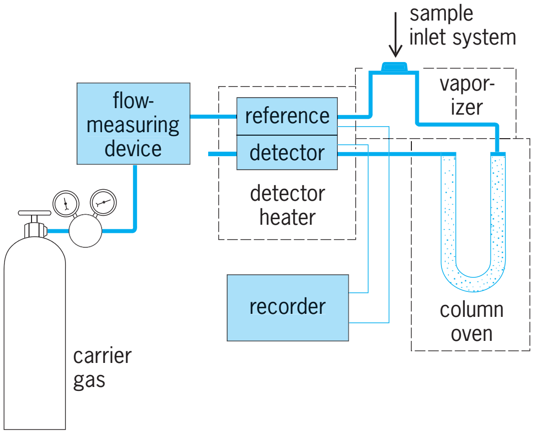


Fig. 2. Basic apparatus used for gas chromatography.

sensitivity of the thermal conductivity detector is a function of the difference in thermal conductivity between the pure carrier gas and that of a mixture of the carrier gas with solute, helium gives high sensitivity. Use of a carrier gas of higher molecular weight will improve column efficiency. Therefore nitrogen or perhaps a gas of even higher molecular weight is preferred to helium if a detector other than thermal conductivity is being used.

A rotometer may be used in the carrier-gas system to give an approximate indication of flow rate. A rotometer consists of a graduated tube with slowly increasing inside diameter and a glass or metal ball that is suspended in the gas flow within the tube at a height dependent upon the flow rate. Since the position of the ball is a function of both the flow rate and the column back pressure when positioned at the column inlet, a rotometer can be used only for rough approximations of flow rate. A soap-bubble flowmeter is used for more accurate measurements.

Sample inlets. These are of two general types depending upon whether the sample is gaseous, liquid, or solid. Liquid samples are generally injected by means of a calibrated hypodermic syringe through a silicon rubber septum into a metal or glass-lined metal injection port, while gaseous samples are introduced by means of a valve and sample loop system. Liquid sample sizes commonly range from 0.5 to 5 microliters. Thus, only very small amounts of material are required for gas chromatographic analysis, 1 microliter being approximately equal to one-fiftieth of a normal-sized drop. Gaseous samples of 0.5–1 milliliter are often used. Injection techniques somewhat similar to those used for liquid samples are used for solids. In order to produce sharp chromatographic peaks with minimum peak overlap, solid and liquid samples must be vaporized rapidly upon injection by maintaining the injection port at a temperature greater than the boiling point of the sample. With some gas chromatographic detectors, the sensitivity is so great that an inlet splitter must be used to reduce the sample amount to one-tenth or one-hundredth of the amount injected. Sample splitters are commonly used with capillary columns, owing to the relatively small amount of liquid phase used with these columns and the much reduced sample capacity which results.

Column. This is the heart of the gas chromatograph, and separation of components on packed columns depends more on the choice of liquid phase than on any other factor. Typically the column is a glass or metal tube of 0.125 or 0.25 in. (3 or 6 mm) in diameter and 4–6 ft (1.2–1.8 m) in length, packed with an inert diatomaceous earth support coated with a nonvolatile liquid to 3–20% by weight. In open tubular or capillary column technology, the support for the thin film of liquid phase is the wall of the capillary itself. Support-coated open tubular columns are also sometimes used, the sample capacity of the columns being increased by the presence of very loosely packed support or by a roughening of the capillary walls.

In choosing a liquid phase for a particular separation, several factors are considered, including solute-solvent interactions (for example, hydrogen bonding), which if present gives special selectivity for a particular solute; temperature limitations of the liquid phase; and the possibility of rare irreversible reactions on the column. Common liquid phases include high-molecular-weight hydrocarbons such as Apiezon L and the natural product, squalane, silicone gum rubbers such as SE-30 and methyl phenyl silicones, poly(ethylene glycol)s such as Carbowax 20M, and certain esters of low volatility such as triethylene glycol succinate. The general rule “like dissolves like” is useful in liquid-phase selection, since it is generally true that polar liquid phases make good solvents for polar solutes while nonpolar liquid phases are often the choice for separation of nonpolar solutes.

In packed column gas chromatography, the liquid phase is adsorbed as a uniform, thin film over the surface of an inert solid support. A good solid support should have the following characteristics: chemical inertness, mechanical strength, relatively large surface area per unit volume, low pressure drop, and thermal stability. Materials with surface areas of 300–1800 ft^2/oz (1–6 m^2/g) are available which fulfill most of these requirements. The most commonly used support in gas chromatography is diatomaceous earth (kieselguhr), which consists of agglomerates of the siliceous skeletons of diatoms treated to produce materials known variously as Chromosorb-P and -W, Celite, C-22 firebrick, Dicalite, Sil-O-Cel, Embacel, and Sterchamol. Chromosorb-G is a diatomaceous earth treated to produce an inert solid, similar in properties to Chromosorb-W but much less friable. To reduce chemical reactivity, the solid support is frequently acid-washed, rinsed, dried, and treated with a silanizing reagent such as trimethylchlorosilane or dimethyldichlorosilane. The supports are generally size-graded to fall in the range 60–80 or 80–100 mesh.

Detectors. The two most popular detectors are the flame ionization detector and the thermal conductivity detector. The flame ionization detector is the more sensitive, with a limit of detection of 0.1–0.001 that of the thermal conductivity detector. The flame ionization detector has a large linear dynamic range, the range in sample size over which the detector responds in a linear manner. In this detector the sample is burned in a small hydrogen flame as it is eluted from the column. Upon combustion of the organic material, charged species are formed in the flame, and these species are collected by a polarized collector electrode. The signal is then fed to an electrometer, a signal being observable for samples as small as 10^{-11} g. Although the detector is sensitive to most carbon compounds, it gives little or no signal for CO_2, CO, CS_2, H_2O, O_2, N_2, NH_3, the rare gases, or oxides of nitrogen and sulfur. The detector has enjoyed widespread use in gas chromatographic analysis and is obviously attractive in pollution studies involving trace organic compounds in air and water owing to its insensitivity toward the air and the water.

The thermal conductivity cell was one of the first gas chromatographic detectors developed. It consists of a metal filament or a thermistor, heated by an electric current from a power supply, and positioned in the gas stream at the end of the column. The filament increases in temperature when organic materials are eluted, since organic compounds have a thermal conductivity less than that of the helium carrier gas, and the heat is therefore conducted away less rapidly.

Wire filaments increase their resistance about 0.4% per degree while thermistors decrease their resistance by about −4% per degree. Thus thermistor thermal conductivity detectors are more sensitive than filaments at low temperatures but become less sensitive at higher detector temperatures. The change in resistance is sensed by making the detector filament a part of a Wheatstone bridge. The advantages of the thermal conductivity detector are that it is sensitive toward all components except the carrier gas being used, and it is nondestructive. It also has a large linear dynamic range, but the limit of detection is less than that of the flame ionization detector.

Other popular gas chromatographic detectors include the flame photometric detector, the electron capture detector, and the thermionic detector. The flame photometric detector may be made selective toward individual metals in organometallic compounds, volatile metal chelates, and volatile metal halides, as well as toward compounds of sulfur and phosphorus. An inexpensive monochromator system or a filter system may be used to isolate the emission wavelength of a particular element, thus making the detector highly selective. The electron capture detector is one of the most sensitive detectors for compounds with electron-capturing properties, such as halogenated compounds and highly conjugated polynuclear aromatic hydrocarbons. It is insensitive to compounds with little or no electron-capturing properties. This detector has proved of particular value in the analysis of pesticides, many of which are chlorinated compounds. If certain alkali metal salts (such as Rb_2SO_4 and CsBr) are placed in the flame of the flame ionization detector and the flow of hydrogen and air is properly adjusted, it is found that the detector becomes particularly sensitive to nitrogen and phosphorus compounds. This orientation is called the alkali flame or thermionic detector.

Quantitative and qualitative analysis. Qualitative information in gas-liquid chromatography is obtained from retention measurements of the peak, and quantitative information is obtained from the area under a peak or, with control of certain variables, from the peak height. Retention time is constant for a given substance on a given column at a given column temperature and carrier-gas flow rate. Thus, a comparison of the retention time of an authentic sample with a component of an unknown mixture gives useful qualitative evidence. Combination of gas chromatography with mass spectrometry provides the ultimate in qualitative information and has been used extensively in research.

The area under a chromatographic peak is measured by means of an electromechanical integrator, an on-line computer system, a planimeter, triangulation approximations (such as multiplication of peak height by the peak width at half height), or sometimes crudely by cutting out the recorded chromatogram and weighing on a balance. Calibration curves may be made by plotting the area or peak height as a function of concentration or amount of sample.

Nonanalytical applications of gas chromatography include measurement of the heat of solution, adsorption, and other thermodynamic quantities; the surface area of powders; the vapor pressure of materials; chemical kinetics; and equilibrium constants. *See* MASS SPECTROMETRY; QUALITATIVE CHEMICAL ANALYSIS.

Important equations. A 4-ft (1.2-m) packed gas chromatography column typically has the equivalent of 600–2000 theoretical plates, *N*. Capillary columns of 1,000,000 plates have been prepared. *N* is a measure of the column efficiency and is given by Eq. (1) where *x* is the distance to the peak maximum

$$N = 16\left(\frac{x}{y}\right)^2 \qquad (1)$$

from the point of injection, and y is the peak width measured at the base. Column efficiency and peak broadening may be evaluated by Eq. (2) where H is

$$H = 2\lambda d_p + \frac{2\gamma D_g}{u} + \frac{8}{\pi^2}\frac{k}{(1+k)^2}\frac{d_l^2}{D_l}u \qquad (2)$$

the average height equivalent to a theoretical plate (= L/N); L is the column length; λ and γ are related to the uniformity in size and shape of the particles in the column; d_p is the average particle diameter; D_g and D_l are the diffusivity of the solute in the gas and liquid phases, respectively; u is the average linear gas velocity; k is the partition ratio, related to the thermodynamic partition coefficient; and d_l is the average thickness of the liquid layer adsorbed on the solid support. This equation shows that optimum column efficiency (smallest value of H) is obtained by using particles of small diameter, uniform in size; a carrier gas of high molecular weight at an optimum velocity of generally 1–4 in./s (2.5–10 cm/s); a nonvolatile liquid phase of low viscosity which will retain the solute well on the column; and a uniform distribution of the liquid phase in a thin layer over the surface of a solid support possessing a large specific surface area. *See* CHROMATOGRAPHY; GAS AND ATMOSPHERE ANALYSIS.

Richard S. Juvet, Jr.

Bibliography. R. L. Grob and E. F. Barry (eds.), *Modern Practice of Gas Chromatography*, 4th ed., 2004; W. Jennings et al., *Analytical Gas Chromatography*, 2d ed., 1997; H. M. McNair and J. M. Miller, *Basic Gas Chromatography*, 1998; W. M. A. Niessen (ed.), *Current Practice of Gas Chromatography-Mass Spectrometry*, 2001.

Gas constant

The universal constant R that appears in the ideal gas law, Eq. (1), where P is the pressure, V the volume, n the amount of substance, and T the thermodynamic (absolute) temperature. The gas constant is universal in that it applies to all gases, providing they are behaving ideally (in the limit of zero pressure). The gas constant is related to the more fundamental Boltzmann constant, k, by Eq. (2), where N_A is the Avogadro constant (the number of entities per mole). The best modern value in SI units is $R = 8.314\,472\,(15)$ J/K · mol, where the number in parentheses represents the uncertainty in the last two digits. *See* BOLTZMANN CONSTANT; GAS.

$$PV = nRT \qquad (1)$$

$$R = N_A k \qquad (2)$$

According to the equipartition principle, at a temperature T, the average molar energy of each quadratic term in the expression for the energy is $(1/2)RT$; as a consequence, the translational contribution to the molar heat capacity of a gas at constant volume is $(3/2)R$; the rotational contribution of a linear molecule is R. *See* KINETIC THEORY OF MATTER.

Largely because R is related to the Boltzmann constant, it appears in a wide variety of contexts, including properties unrelated to gases. Thus, it occurs in Boltzmann's formula for the molar entropy of any substance, Eq. (3), where W is the number of arrangements of the system that are consistent with the same energy; and in the Nernst equation for the potential of an electrochemical cell, Eq. (4), where E° is a standard potential, F is the Faraday constant, and Q is a function of the composition of the cell. The gas constant also appears in the Boltzmann distribution for the population of energy levels when the energy of a level is expressed as a molar quantity. *See* BOLTZMANN STATISTICS; ELECTRODE POTENTIAL; ENTROPY.

$$S = R \ln W \qquad (3)$$

$$E = E^\circ - (RT/nF) \ln Q \qquad (4)$$

P. W. Atkins

Bibliography. P. W. Atkins, *Physical Chemistry*, 6th ed., 1998.

Gas discharge

A system made up of a gas, electrodes, and an enclosing wall in which an electric current is carried by charged particles in response to an electric field, the gradient of the electric potential, or the voltage between two surfaces. The gas discharge is manifested in a variety of modes (including Townsend, glow, arc, and corona discharges) depending on parameters such as the gas composition and density, the external circuit or source of the voltage, electrode geometry, and electrode material. A gas discharge can also be inductively coupled to an alternating-current (ac) circuit, eliminating the need for electrodes. Gas discharges are useful both as tools to study the physics existing under various conditions and in technological applications such as in the lighting industry and in electrically excited gas lasers. *See* LASER; VAPOR LAMP.

The study of gas discharges encompasses a broad range of topics in physics and engineering, only a few of which are mentioned here. New applications in gas insulation, in high-power electrical switching, and in materials reclamation and processing will assure a continuing effort to better understand all aspects of gas discharges. *See* GAS TUBE.

Sources of electrons. Electrons, rather than ions, are the main current carriers in gas discharges because their mass is smaller and their mobility is correspondingly much higher than that of ions. Electrons are produced by ionization of the gas itself, or they originate at the electrodes present in the system. Gas ionization can be accomplished in several ways, including electron impact ionization, photoionization, and associative ionization. Bombardment by photons, energetic ions or electrons, and excited neutral particles can cause secondary emission from the electrodes. A high-energy-per-unit electrode surface area can induce thermionic or field emission of electrons. Each means of producing electrons leads to a different response of the gas discharge as a circuit element. *See* ELECTRON EMISSION; FIELD EMISSION; IONIZATION; PHOTOEMISSION; SECONDARY EMISSION; THERMIONIC EMISSION.

Townsend discharge. A nonself-sustained discharge is one in which an agent external to the circuit is necessary to replenish the electrons lost from the system in order to maintain a flow of current; an example is a Townsend discharge. A Townsend discharge, considered here in a parallel plate electrode geometry, is a low-current (less than 10^{-5} A) discharge in which the primary source of current carriers is, for example, photoemission resulting from illumination of the cathode. The electrons gain energy from the electric field between the electrodes and lose energy to the gas through elastic and inelastic collisions with the neutral gas particles. Except in regions very near the electrodes, the energy gained by the electrons from the field balances that lost in collisions, and the discharge characteristics can be parametrized simply by the neutral gas composition and by the ratio of the field strength E to the neutral gas density n. Townsend discharges have been widely used to study the physics of the interaction of electrons with neutral gas particles because external control of E/n provides a convenient means for adjusting the electron energy. *See* TOWNSEND DISCHARGE.

Gas breakdown. For E/n greater than about 10^{-19} V · m^2 (100 townsends) some of the electrons have sufficient energy to produce additional ionization in collisions with the neutral gas particles. The succeeding generation of electrons may also produce further ionization, and if the volume rate of ionization exceeds the volume losses due to diffusion and attachment, an exponential increase in the electron density results. This cascade is termed an electron avalanche. If the avalanche grows to the point where the gas loses its insulating properties and becomes highly conducting, gas breakdown is said to have occurred. The luminosity accompanying gas breakdown is due to electron impact excitation of the gas particles which then radiate. Lightning is a familiar example of the effect. *See* ELECTRICAL BREAKDOWN.

In determining the minimum voltage required for a self-sustained discharge, electron emission from the electrodes must be considered as well as the cascade ionization; that is, electrons originating from secondary emission at the electrodes must be considered in determining the minimum applied voltage for which the electron sources in the discharge balance the losses. The Paschen curve shows that this minimum voltage, or sparking potential, is a function of the product of the gas pressure p and the electrode separation d. The sparking potential at a given value of pd depends on gas composition, electrode material and geometry, and applied field frequency, but a minimum in the sparking potential always occurs at a finite value of pd. *See* BREAKDOWN POTENTIAL; ELECTRIC SPARK.

Glow discharge. At pressures lower than about 100 torr (13 kilopascals), but depending on the gas composition, a succession of overlapping avalanches between two electrodes leads to the very stable glow discharge, where space-charge fields adjust themselves so that the volume production of ionization, and the resulting secondary emission from the electrodes, balances the loss of the current carriers. The glow discharge is characterized by moderate current (10^{-5} to 10^{-2} A) and voltage (on the order of 100 V), a particular field distribution due to the space-charge distortion of the applied field, and a well-defined luminosity pattern between the electrodes. In addition, the average electron energy is considerably higher than the average kinetic or internal energies of the neutral particles. In the normal glow discharge, the discharge voltage and the current density are nearly constant. An increase in current is achieved by reducing the external circuit impedance or by increasing the source voltage, which causes an increase in the electron-emitting area on the cathode surface. Once the glow covers the cathode surface, an additional increase in current is associated with an increase in the discharge voltage and in the current density. This regime is termed the abnormal glow. *See* GLOW DISCHARGE.

Arc. By lowering the circuit resistance, or increasing the source voltage, a further increase in current can be realized. Past a certain current level, however, this increase is accompanied by an abrupt change in the discharge characteristics, and this high-current discharge mode is termed an arc. An arc is distinguished operationally from a glow by the large currents and the small voltage drop. A more fundamental distinction can be made by considering the sources of the ionization. A low-pressure arc is sustained by thermionic or field emission of electrons from the electrode surfaces; secondary electron emission and electron impact ionization are still present but are not very important as electron sources. At higher pressures, volume sources of electrons dominate, but the sources involve multistep, cumulative ionization through excited states, rather than the electron impact ionization predominantly from ground-state neutral particles, which is characteristic of most glow discharges. The range of applications of arcs is enormous. *See* ARC DISCHARGE; ARC LAMP; ARC WELDING.

Glow-to-arc transition. At pressures higher than about 100 torr (13 kPa), depending on the gas composition, the glow discharge is not stable and has a tendency to constrict to form an arc. The details of the mechanisms for the glow-to-arc transition are not clearly established, but involve changes in the kinetic and internal energy of the gas due to the large amounts of energy deposited in the gas. In a high-pressure arc, the gas is hot and highly excited, and collisions between electrons or neutral particles and excited gas particles produce most of the ionization; that is, the system approaches local thermodynamic equilibrium described by temperatures on the order of 10^3 to 10^4 K (2×10^3 to 2×10^4 °F). Diffusion at the lower pressures tends to dissipate the energy in the gas and to stabilize the discharge in the glow mode, but the lower diffusion at higher pressures leads to nonuniform gas heating and excitation in the center of the discharge and the associated highly conducting arc. Any other nonuniformities in the system, either in the gas or on the electrodes, can also provide a seed for arc formation at high pressures. Arcs are undesirable in applications that depend on uniform volume excitation of the gas such as electrically excited gas lasers.

Corona discharge. This discussion has implicitly focused on parallel plate electrode geometries. Discharges in irregular geometries where the applied field confines the current density to localized regions are called corona discharges. These relatively low-current, high-voltage discharges occur in Townsend or glow discharges with point-plane electrode configurations or around conductors with a small radius of curvature, and are often intermittent. Corona discharges can represent significant power losses in high-voltage transmission lines. *See* CORONA DISCHARGE; ELECTRICAL CONDUCTION IN GASES.

Leanne C. Pitchford

Bibliography. R. W. Crompton et al. (eds.), *Gaseous Electronics and Its Applications*, 1991; M. N. Hirsh and H. J. Oskam (eds.), *Gaseous Electronics*, vol. 1: *Electrical Discharges*, 1978; E. E. Kunhardt and L. H. Luessen (eds.), *Electrical Breakdown and Discharges in Gases*, 2 vols., 1983; J. M. Meek and J. D. Craggs (eds.), *Electrical Breakdown in Gases*, 1978; Y. P. Raizer and J. E. Allen, *Gas Discharge Physics*, 1997.

Gas dynamics

The study of gases in motion. In general, matter exists in any of three states: solid, liquid, or gas. Liquids are incompressible under normal conditions; water is a typical example. In contrast, gases are compressible fluids; that is, their density varies depending on the pressure and temperature. The air surrounding a high-speed aircraft is an example. *See* GAS; LIQUID.

Scope of subject. Gas dynamics can be classified in a variety of ways. The following discussion deals with gases as a continuum. The structure of gases on the particle level is called rarefied gas dynamics which is discussed in the last section of the article. *See* AEROTHERMODYNAMICS; COMPRESSIBLE FLOW; FLUID FLOW.

Fundamental relations. Gases in motion are subject to certain fundamental laws. These are the laws of the conservation of mass, momentum, and energy. They apply as long as the velocity does not approach the speed of light, when relativistic effects must be taken into consideration. This regime is outside the scope of this discussion. The conservation laws are represented by the continuity-of-mass equation, the conservation-of-momentum equation, and the conservation-of-energy equation. In other words, the fundamental conservation principles of mechanics and dynamics constitute the theoretical basis of gas dynamics. In the case of the dynamics of incompressible fluids, it is usually sufficient to satisfy only the laws of conservation of mass and momentum. This distinction constitutes the fundamental difference between high-speed gas dynamics and hydrodynamics. If irreversibilities are involved, a fourth equation called the entropy balance equation may be considered. Whereas mass, momentum, and energy are conserved, the entropy is not. Real problems are irreversible; that is, losses such as friction are involved. However, as a first approximation such effects are generally not considered. *See* CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; CONSERVATION OF MASS; CONSERVATION OF MOMENTUM.

The mass, momentum, and energy equations are higher-order, nonlinear equations that have no general solution, and only a limited number of special cases can be solved. Another approach is to resort to numerical solutions using high-speed digital computers. While this approach has proven to be very useful, it limits the degree to which flow phenomena can be conceptualized. Accordingly, it is frequently permissible to write the equations in one-dimensional form. By one-dimensional flow is meant that the properties of gas such as its velocity and density are assumed to be constant in the direction perpendicular to the direction of the gas flow. Generally, the one-dimensional approach gives excellent insights into understanding the physical behavior of the flow. It is also very useful in setting up the computer algorithm for numerical solutions. *See* COMPUTATIONAL FLUID DYNAMICS; DIFFERENTIAL EQUATION.

Continuity equation. One other matter must be considered, namely whether the flow is steady or unsteady. In steady flow, the flow characteristics do not vary with time, whereas unsteady flow implies that the flow assumes different configurations over time. Thus, unsteady flow is broader in scope. In this case the continuity equations for conservation of mass may be written as Eq. (1). In this equation, the first

$$\frac{\partial(\rho V)}{\partial x} + \frac{\partial \rho}{\partial t} = 0 \tag{1}$$

term defines the mass-flow changes with respect to the space coordinates, whereas the second term indicates the changes with time. Here, ∂ is the partial differential operator; x denotes the space coordinate, in this case the direction of flow; ρ is the gas density; V is the gas velocity; and t is the time.

If the flow is steady, there is no time-dependent term, and hence the continuity equation can be written in integrated form as Eq. (2), where A denotes

$$\rho V A = \text{constant} \tag{2}$$

the area in the direction perpendicular to the flow direction.

Momentum equation. The momentum equation is the mathematical representation of the law of conservation of momentum. It is a statement of the forces acting on the gas. Different types of forces must be recognized. Body forces, such as gravitation and electromagnetic forces, act at a distance. The so-called surface forces may assume different forms, such as normal stresses and viscosity. The simplest form of the momentum equation relevant to gas dynamics, which applies to steady flow without friction, is Eq. (3). This equation demonstrates that, where the

$$-dp = \rho V dV \tag{3}$$

pressure increases along the flow streamlines, the velocity decreases. For gases at very low speed (nearly constant density), Eq. (3) may be integrated to give the familiar Bernoulli equation (4), where the con-

$$\frac{p}{\rho} + \frac{V^2}{2} = \text{constant} \tag{4}$$

stant can be reduced by friction. *See* BERNOULLI'S THEOREM.

Energy equation. The energy equation expresses the first law of thermodynamics and accounts for the changes in energy as the gas moves about its path.

It can also take into consideration energy exchanges between the gas and its environment, such as radiation. The simplest form of the energy equation for compressible gas flow, which applies along streamlines in the absence of heat transfer, is Eq. (5), where

$$c_pT + \frac{V^2}{2} = \text{constant} \tag{5}$$

c_p denotes the specific heat at constant pressure and T is the gas temperature. Equation (5) shows that where temperature rises the speed decreases and vice versa. The constant in Eq. (5) can be increased by heat transfer to the gas.

The above review intentionally presented the governing equations of gas dynamics in their elementary forms. Incorporating all possible effects requires elaborate discussions that can be found in textbooks.

Speed of sound. The speed of sound or the acoustic velocity is a very important term in gas dynamics because it serves as a criterion to identify flow regimes. Being able to do so is crucial because the designer must know the conditions that the gas will generate or, conversely, experience. The acoustic velocity (a) is defined by Eq. (6), where γ is the specific-heat

$$a = \sqrt{\gamma RT} \tag{6}$$

ratio and R is the gas constant.

In prescribing flow regimes, the flow velocity of the gas is compared with the acoustic velocity. This ratio, called the Mach number (M), is defined by Eq. (7). Using the Mach number the following flow

$$M = \frac{V}{a} \tag{7}$$

regimes are described:

$M < 1$	subsonic flow
$M = 1$	sonic flow
$0.9 < M < 1.1$	transonic flow
$M > 1$	supersonic flow
$M > 5$	hypersonic flow

High-speed aircraft are categorized by the Mach number. The Mach number is significant otherwise. Squaring both sides of Eq. (7) yields Eq. (8), which

$$M^2 = \frac{V^2}{a^2} \tag{8}$$

can be recognized as the ratio of the kinetic energy to the thermal energy. This concept is very helpful in the manipulation of gas-dynamic equations because gas dynamics is really a combination of dynamics and thermodynamics. *See* MACH NUMBER; SOUND; THERMODYNAMIC PRINCIPLES.

Wave phenomena. Consider a pulsating pressure source moving with velocity V. This source starts pulsating at time $t = 0$. In a time interval t_1, the source travels a distance Vt_1, while the signal reaches the surface of a sphere of radius at_1 (the surface of this sphere constitutes the wavefront at time t_1). At a later time t_2, the source point is at Vt_2, while the signal front is at at_2. Thus, as long as $V/a = M < 1$, the source point is always inside the outermost wavefront.

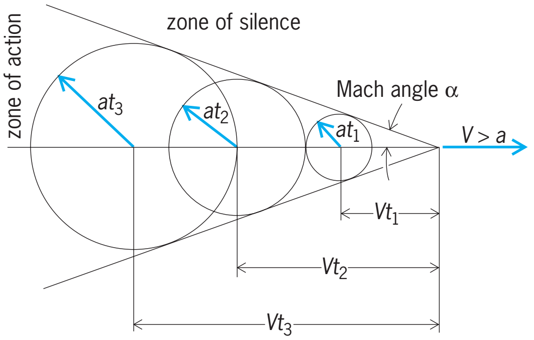


Fig. 1. Rule of forbidden signals from a point source moving at supersonic velocity. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

Still another use of the term wave refers to a wave envelope. If the point source moves so fast that $V > a$, the wavefront spheres will no longer contain the source (**Fig. 1**). The envelope to this family of spheres is a cone, known as the Mach cone, and the Mach angle is such that $\sin \alpha = 1/M$.

The Mach line constitutes a demarcation (Fig. 1). The fluid outside the Mach line will receive no signal from the source. T. von Kármán has appropriately called this phenomenon the rule of forbidden signals and designated the region ahead of the Mach line the zone of silence and the region inside the Mach line the zone of action. *See* HYPERSONIC FLIGHT; SUPERSONIC FLIGHT; TRANSONIC FLIGHT.

Shock waves. In the same manner in which a Mach wave is the envelope of infinitesimal disturbances, a shock wave is the envelope of finite disturbances. The steady conditions on either side of a standing shock wave can be obtained by applying the conservation laws expressed by Eqs. (9), (10), and (11).

$$\rho_1 V_1 = \rho_2 V_2 \qquad \text{(continuity)} \tag{9}$$

$$p_1 + \rho_1 V_1^2 = p_2 + \rho_2 V_2^2 \qquad \text{(momentum)} \tag{10}$$

$$h_1 + \frac{V_1^2}{2} = h_2 + \frac{V_2^2}{2} \qquad \text{(energy)} \tag{11}$$

In Eq. (11), $h = u + p/\rho$, where u is the internal energy. The quantity h is termed enthalpy, and for a perfect gas it is given by c_pT. By a simple rearrangement of these equations, the approximate expressions are obtained for the shock Mach number in the case of a perfect gas, Eq. (12) for weak shocks and Eq. (13) for strong shocks.

$$M_s \approx 1 + \frac{\gamma + 1}{4\gamma} \frac{p_2 - p_1}{p_1} \tag{12}$$

$$M_s \approx \left(\frac{\gamma + 1}{2\gamma} \frac{p_2}{p_1} \right)^{1/2} \tag{13}$$

In a sound wave, $p_2 - p_1 \simeq 0$ and, therefore, $M_s = 1$; for $p_2/p_1 = 4$, the shock speed is roughly twice the

	M_1	M_2	p_2/p_1	V_2/V_1	ρ_2/ρ_1	T_2/T_1
detonation	>1	≶1	>1	<1	>1	>1
deflagration	<1	<1	<1	>1	<1	>1

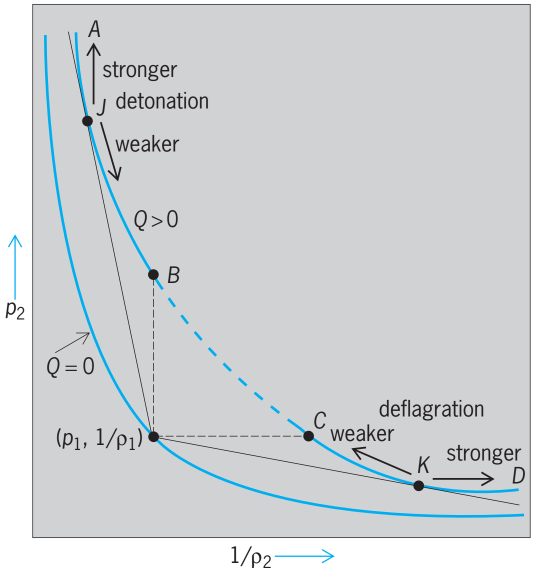


Fig. 2. Hugoniot diagram and Chapman-Jouget conditions. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

speed of sound, as shown in Eq. (13). *See* ENTHALPY; SHOCK WAVE.

Detonation and deflagration waves. Other interesting gas-dynamic waves are characterized by the same continuity and momentum equations. The energy equation, however, is modified to include a term which accounts for chemical heat release (Q); such waves are either detonations or deflagrations.

Eliminating the kinetic energy from the energy equation by the use of the continuity equation and the momentum equation yields Eq. (14).

$$h_1 - h_2 + Q = \frac{1}{2}(p_1 - p_2)\left(\frac{1}{\rho_2} + \frac{1}{\rho_1}\right) \qquad (14)$$

For a given Q, zero or nonzero, and given p_1 and ρ_1 (which through the equation of state gives h_1) and one additional variable behind the wave, for example, V_2 or h_2, the locus of all possible combinations of ρ_2 and p_2 can be plotted on a so-called Hugoniot diagram (**Fig. 2**). The lines OJ and OK are tangent to the Hugoniot curve; O is the point p_1, $1/\rho_1$. Points J and K separate strong and weak waves. Flows corresponding to J and K are characterized by the fact that all the thermodynamic and fluid-mechanic variables have an extremum. Transitions from B to C or vice versa involve a decrease in entropy and are therefore forbidden. *See* ENTROPY.

Mach number functions. In many applications it is reasonable to assume that the gas is perfect—both thermally and calorically—and that the flow is adiabatic. It then becomes very useful to express all the dependent variables in terms of the Mach number.

Starting with the energy equation, a stagnation temperature T_0 is defined by expression (15).

$$\frac{V^2}{2} + c_pT = c_pT_0 \qquad (15)$$

Physically, T_0 represents the temperature the gas would have if all its kinetic energy were transformed into thermodynamic enthalpy. The stagnation enthalpy can be measured by a thermometer immersed in a gas stream.

From Eq. (14), a simple derivation yields Eq. (16),

$$\frac{T_0}{T} = 1 + \frac{\gamma - 1}{2}M^2 \qquad (16)$$

where T is now called the static temperature. This temperature is measured with a thermometer at rest with respect to the gas. The stagnation temperature is constant in any adiabatic flow, even through a shock, and thus provides an excellent reference parameter. By using the isentropic relation $p = \rho^\gamma$ and the perfect gas law, Eq. (17), a reference stagnation

$$P = \rho RT \qquad (17)$$

pressure and density can be defined by means of Eqs. (18) and (19).

$$\frac{p_0}{p}\left(1 + \frac{\gamma - 1}{2}M^2\right)^{\gamma/(\gamma-1)} \qquad (18)$$

$$\frac{\rho_0}{\rho}\left(1 + \frac{\gamma - 1}{2}M^2\right)^{1/(\gamma-1)} \qquad (19)$$

When $M^2 \ll 1$, Eq. (18) may be expanded by the binomial theorem to obtain Eq. (20).

$$p_0 = p + \frac{1}{2}\rho V^2\left(1 + \frac{M^2}{4} + \cdots\right) \qquad (20)$$

The deviation from Bernoulli's equation, Eq. (4), due to compressibility at $M = 0.5$ is only 6%.

Flows can be classified as internal flow and external flow. Internal flow refers to the cases where the gas is constrained by a duct of some sort. Characteristically external flow is flow over an airplane or missile.

Internal one-dimensional flow. Internal flows are conveniently characterized by (1) the shape of the duct and its variation, (2) the heat transfer through the walls of the duct and internal heat sources, and (3) frictional effects. By varying one of these characteristics at a time, the essential features of internal flow can be discussed most simply.

Variable area flow. A device to accelerate the flow of a gas or liquid is termed a nozzle. In most engineering applications the contour of the nozzle is first converging and then diverging; it thus has a minimum cross section, called a throat. *See* NOZZLE.

For isentropic flow in a convergent-divergent nozzle in which the flow is supersonic in the divergent section, the velocity at the throat is sonic, that is, $M^* = 1$. The throat pressure is then said to be

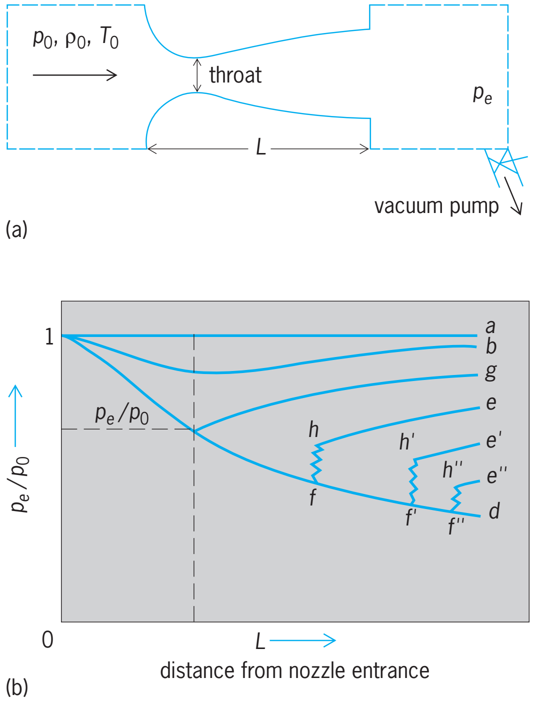


Fig. 3. Convergent-divergent nozzle between two reservoirs. (*a*) Diagram of nozzle. (*b*) Pressure distribution in nozzle for given reservoir pressure (p_0) and various values of exit pressure (p_e) (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*).

critical p^* and is given by Eq. (21). Velocity and pressure are related in this case by Eq. (22).

$$p^* = p_0\left(\frac{2}{\gamma+1}\right)^{\gamma/(\gamma-1)} \tag{21}$$

$$V = \frac{2\gamma R}{\gamma-1}T_0[1-(p/p_0)^{(\gamma-1)/\gamma}] \tag{22}$$

A convergent-divergent deLaval nozzle can be inserted between two reservoirs (**Fig.** 3*a*), and the pressure can be plotted as a function of the distance from the nozzle entrance (Fig. 3*b*). There will be no flow if the ratio of exit pressure to reservoir pressure is $p_e/p_0 = 1$, as represented by curve *a*. If p_e is reduced so that it is slightly less than the entrance pressure, the nozzle will act like a conventional venturi, as represented by curve *b*. For this case, the flow is always subsonic and resembles incompressible flow. When the exit pressure is reduced further, the critical pressure can be reached at the throat, as curve *g* shows. In this case, the velocity is sonic at the throat but is never supersonic within the nozzle, even though the pressure at the throat corresponds to the critical. The minimum pressure which can exist in the nozzle outlet is depicted by point *d*, for which the pressure at the throat will be the critical. Here the velocity in the converging section is subsonic, in the diverging section it is supersonic, and at the throat it is sonic. For the range of exit pressures from p_d to p_g, the rate-of-flow remains constant for a given reservoir pressure p_0. The flow rate reaches a maximum value and remains there over this wide range of exit pressures.

Even in the absence of friction, isentropic flow can exist only for the range of exhaust pressures from p_a to p_g and at the pressure reached along curve *d*, but not at intermediate pressures. The pressures p_g and p_d are the significant design pressures for a given nozzle. For exhaust pressures in the range between p_d and p_g, shocks will occur in the nozzle, raising the pressure from *f* to *h* (or *f*′ to *h*′), followed by a pressure rise after the shock points to an exit pressure such as p_e. If p_e is less than p_d, the jet leaving the nozzle is said to be underexpanded and will drop in pressure after leaving the mouth of the nozzle. The velocity V_1 in front of a shock is supersonic, and the velocity V_2 behind a normal shock (in contrast to an oblique shock) is always subsonic. For a normal shock $V_1V_2/a^{*2} = 1$, where a^* is the critical speed of sound corresponding to $M = 1$. Thus for $V_1 > V_2$, $V_1/a^* > 1$ and therefore $V_2/a^* < 1$. *See* ISENTROPIC FLOW.

Diabatic flow. Heat exchangers and combustion chambers are devices in which heat transfer occurs. The equations describing nonadiabatic or diabatic processes are complicated; consequently, certain limiting assumptions are usually required to make possible analytical solutions of the equations.

These assumptions are that (1) the flow takes place in a constant-area section, (2) there is no friction, (3) the gas is perfect and has constant specific heats, (4) the composition of the gas does not change, (5) there are no devices in the system which deliver or receive mechanical work, and (6) the flow is steady.

Equations which conform to these requirements are called Rayleigh equations, and the associated flow is designated as Rayleigh flow. Designating by $Q_{1\rightarrow 2}$ the quantity of heat introduced between stations 1 and 2, Eq. (23) is obtained for the energy equation.

$$\begin{aligned} Q_{1\rightarrow 2} &= c_p(T_2 - T_1) + \frac{V_2^2 - V_1^2}{2} \\ &= h_2 - h_1 + \frac{V_2^2 - V_1^2}{2} \end{aligned} \tag{23}$$

If the stagnation enthalpy is introduced, then $Q_{1\rightarrow 2}$ is given by Eq. (24), which can be expressed in terms of stagnation temperatures by Eq. (25).

$$Q_{1\rightarrow 2} = h_{02} - h_{01} \tag{24}$$

$$Q_{1\rightarrow 2} = c_p(T_{02} - T_{01}) \tag{25}$$

Because $Q_{1\rightarrow 2} \neq 0$ for diabatic flow and because $c_p > 0$ always, it follows that $T_{02} \neq T_{01}$. This inequality states that in diabatic flow the stagnation temperature is not solely determined by the reservoir conditions, as is the case with adiabatic flow. Heating raises the stagnation temperature; cooling lowers it. *See* ADIABATIC PROCESS.

The locus of points of properties during a constant-area, frictionless flow with heat exchange is called the Rayleigh line. By definition, along the Rayleigh line the continuity equation and the

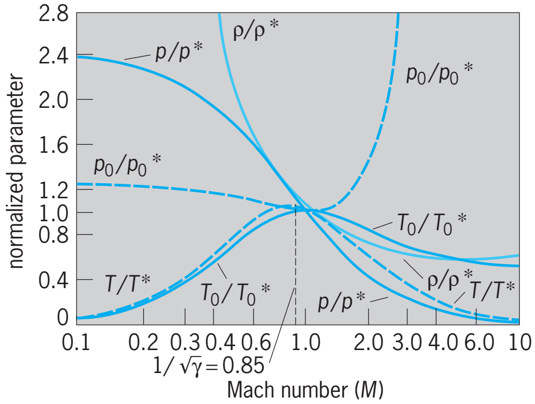


Fig. 4. Diabatic flow parameters for specific heat ratio of 1.4. Asterisk indicates value of parameter when $M = 1$. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

momentum equation must apply. Equation (9) applies to steady flow in a constant-area duct. Mass velocity by definition is $G = \rho V$ and, from Eq. (10), the momentum relation is $p + \rho V^2 = C$, where C is a constant. Consequently, Eq. (26) is obtained, which

$$p + \frac{G^2}{\rho} = C \tag{26}$$

is one of the many Rayleigh-line equations.

The variations of pressure, temperature, and density with Mach number may be plotted for Rayleigh flow (**Fig. 4**). The fact that the curve for T_0/T_0^* reaches a maximum at a Mach number of unity indicates that it is impossible to pass from one flow domain into the other by the same heat-transfer process. Thus, if heat is added to a subsonic flow, the flow can be accelerated only until its Mach number becomes unity. Further addition of heat will not further accelerate the gas but will result in choking of the flow. As a consequence, the flow must readjust itself, which it will do by lowering its initial Mach number (**Table 1**). *See* CHOKED FLOW.

Flow with friction. In long pipes the effects of friction may result in a significant pressure drop. Over a length dx, this pressure drop dp is given by the Fanning equation (27), where f is a friction factor that

$$dp = -f\frac{\rho V^2}{2r_h}dx \tag{27}$$

must be determined experimentally and r_h is the hydraulic radius of the pipe. To solve friction-flow problems analytically, certain simplifying assumptions are made and the resulting hypothetical flow is called Fanno flow. The Fanno flow assumptions are the same as those for Rayleigh flow except that the assumption that there is no friction is replaced by the requirement that the flow be adiabatic. Numerous Fanno flow equations may be written by combining the energy and the continuity equations in accordance with these assumptions. The variation of properties during Fanno flow may be plotted (**Fig. 5**), and the trends of the most important properties during subsonic and supersonic flow may be summarized in tabular form (**Table 2**).

The Fanno and Rayleigh lines may be plotted for the same constant mass velocity $G = \rho V$ (**Fig. 6**). The Rayleigh and Fanno lines have two points of intersection, denoted by a and b; a normal shock connects these two points. The flow through a shock wave is irreversible; thus, associated with it is an increase in entropy. Point b, thus, always lies to the right of point a.

There is another interesting point about the Fanno curve. If frictional flow continues along the subsonic portion of the Fanno line, the Mach number tends to increase toward unity, whereas if it continues along the supersonic portion, the Mach number decreases toward unity. As in the case of Rayleigh flow, it is impossible, by virtue of the second law of thermodynamics, to pass from one flow regime to the other (subsonic into supersonic or conversely) unless the mass velocity is readjusted.

External flow. Boundary layers and wakes are the centers of interest in external flows. Here the effects of compressibility are substantially more difficult to analyze than in internal flows, if for no other reason than the inapplicability of a one-dimensional approach.

Ballistic missiles and space vehicles enter the Earth's atmosphere with velocities typically of 4 mi/s (6–7 km/s). The corresponding Mach number, depending on the altitude, is of the order of 20. The energy that maintains the bow shock and the work done to overcome the viscous shear stresses, Φ, reduce the kinetic energy of the vehicle, and it decelerates. This loss of kinetic energy, which is the work of the drag forces, reappears in part as the increased enthalpy and temperature of the fluid near the vehicle surface. The rate of heat transfer driven by the large enthalpy gradient $(h_g - h_s)/\Delta$, where h_g is the gas enthalpy at some suitably defined distance Δ from the surface and h_s is the enthalpy right at the

TABLE 1. Variation of flow properties for Rayleigh flow

	Heating		Cooling	
Property	$M > 1$	$M < 1$	$M > 1$	$M < 1$
T_0	Increases	Increases	Decreases	Decreases
p	Increases	Decreases	Decreases	Increases
p_0	Decreases	Decreases	Increases	Increases
V	Decreases	Increases	Increases	Decreases
T	Increases	Increases when $M < 1/\sqrt{\gamma}$ Decreases when $M > 1/\sqrt{\gamma}$	Decreases	Decreases when $M < 1/\sqrt{\gamma}$ Decreases when $M > 1/\sqrt{\gamma}$

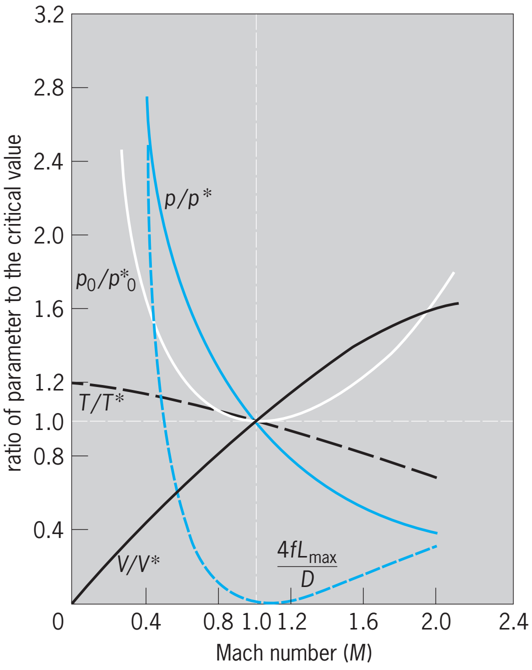


Fig. 5. Functions for constant-area flow with friction (γ = 1.4). Asterisk indicates value of parameters when $M = 1$. f = friction factor; L_{max} = maximum length of pipe without occurrence of choking; D = pipe diameter. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

TABLE 2. Fanno flow phenomena

Property	Initial flow is subsonic	Initial flow is supersonic
M	Increases	Decreases
V	Increases	Decreases
p	Decreases	Increases
T	Decreases	Increases
ρ	Decreases	Increases

surface, is so large that special protection must be afforded the vehicle. To this purpose, the vehicle can be covered with a material designed to char, melt, or gasify and in so doing absorb much of the heat that would otherwise penetrate into the structure. This "ablation" process, as it is called, introduces large amounts of material (some of it chemically active, some of it ionized, and some of it radiating) into what ordinarily would be called a boundary layer. Unfortunately, many of the assumptions that make it possible to introduce the boundary-layer simplifications are violated here. Very complex computer programs have been developed, however, that yield reasonably accurate estimates of these effects. *See* ATMOSPHERIC ENTRY; BOUNDARY-LAYER FLOW.

Joshua Menkes; Ali B. Cambel

Rarefied gas dynamics. Rarefied gas dynamics is that branch of gas dynamics dealing with the flow of gases under conditions where the molecular mean free path is not negligibly small compared to some characteristic dimension of the flow field. Rarefied flows occur when the gas density is extremely low, as in the cases of vacuum systems and high-altitude flight, but also when gases are at normal densities if the characteristic dimension is sufficiently small, as in the case of very small particles suspended in the atmosphere.

The dimensionless parameter which describes the degree of rarefaction existing in a flow is the Knudsen number, $\mathrm{Kn} = \lambda/L$, defined as the ratio of the mean free path λ to some characteristic dimension L of the flow field. Depending on the situation, L might be chosen, for example, as the diameter of a duct in a vacuum system, the wavelength of a high-frequency sound wave, the diameter of a suspended submicrometer-size particle, the length of a high-altitude rocket, or the thickness of a boundary layer or a shock wave. The mean free path λ, which is the average distance traveled by a gas molecule between successive collisions with other molecules, is equal to the molecular mean speed, given by Eq. (28), divided by the collision frequency ν_c: thus, Eq. (29)

$$\bar{C} = \sqrt{\frac{8}{\pi}RT} \tag{28}$$

$$\lambda = \frac{\bar{C}}{\nu_c} \tag{29}$$

is satisfied. However, it is often more convenient in evaluating the Knudsen number to use the viscosity-based mean free path given by Eq. (30), where ν is

$$\lambda \simeq \frac{2\nu}{\bar{C}} \tag{30}$$

the kinematic viscosity. *See* VISCOSITY.

Flow regimes. It is convenient to divide rarefied flows into three flow regimes, according to the range of values of the appropriate Knudsen numbers. The regime of highly rarefied flow, which obtains for Kn much greater than 1 (typically greater than 10), is called collisionless or free-molecule flow, while the regime of slight rarefaction, where Kn is much less than 1 (typically less than 0.1), is called slip flow. Flows at Knudsen numbers intermediate to these limiting values are termed transition flows. The phenomena and methods of analysis associated with the

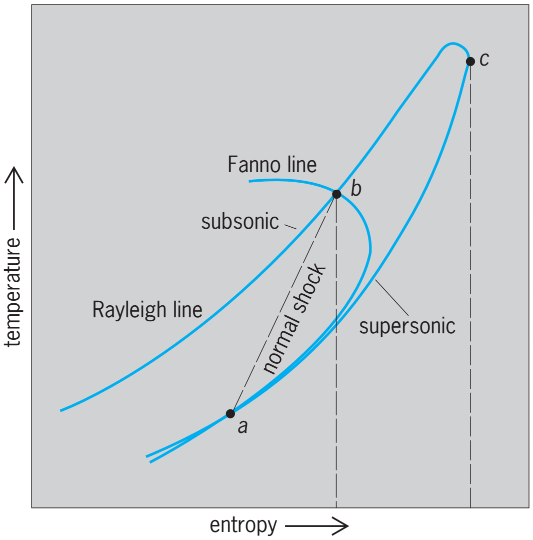


Fig. 6. Rayleigh and Fanno lines. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

three regimes are in general quite dissimilar, so the classifications are helpful.

Collisionless flow. Collisionless flows are ones for which intermolecular encounters are very rare. Thus, in the case of a vehicle in high-altitude flight, such as a satellite, molecules impinging on a surface will travel, after reflection, very far before colliding with other molecules, with the consequence that the incoming flux of molecules is unperturbed by the reflected flux. For many applications, such as high-altitude aerodynamics, it can be assumed that the molecular velocity distribution f_i of the incident molecules is a drifting maxwellian given by Eq. (31),

$$f_i = n\left(\frac{m}{2\pi kT}\right)^{3/2} e^{-mc^2/2kT} \tag{31}$$

where m is the molecular mass, n the number density, k Boltzmann's constant, T the free stream temperature, and the peculiar molecular velocity $\mathbf{C}$ is the difference between the absolute molecular velocity $\boldsymbol{\xi}$ and the mean gas velocity $\mathbf{V}$, which in the case of an aerodynamic body in steady flight would be its flight velocity. With this assumption, the incident fluxes of mass, momentum, and energy on a unit area of a body can be evaluated readily in terms of appropriate moments of f_i. *See* BOLTZMANN STATISTICS; KINETIC THEORY OF MATTER; STATISTICAL MECHANICS.

The net fluxes to the surface element are determined by the nature of the velocity distribution f_r of the reflected molecules. Unfortunately, f_r is generally not known, and recourse must be had to empirical models of the molecular reflection process. One limiting case is specular reflection, in which the molecular velocity component normal to the surface is reversed, and the tangential velocity component remains unchanged. However, this is not characteristic of real gas-solid surface interactions. Another limiting case is diffuse reflection, for which the velocities of the molecules after reflection are independent of their incident velocities, and are in a half-range maxwellian distribution at the wall temperature T_W. The degree to which an actual surface behaves in a diffusive manner is customarily measured by three surface interaction parameters, or accommodation coefficients.

Calculation of the incident and reflected fluxes of energy and momentum leads to evaluation of the local pressure, shear stress, and energy transport to a surface element of a body. Then by integration over the entire body surface, overall lift, drag, and heat-transfer characteristics are obtained. The calculations are straightforward for simple convex shapes such as flat plates, spheres, and cylinders.

The analysis of collisionless flows becomes more complex if molecules can undergo multiple surface interactions, as in flows in ducts and over concave surfaces.

Slip flow. This is the regime of slight rarefaction, and is manifested initially as an alteration of the boundary conditions associated with the basic flow equations, the Navier-Stokes equations. The typical length scale L defining the Knudsen number is the boundary-layer thickness δ, although it could be the diameter of a pipe in the case of an internal flow. Boundary layers in the slip flow regime are usually laminar, but may exhibit strong compressibility and heat-transfer effects in high-speed flows. Slip relations are applicable only under conditions of small rarefaction, typically $\lambda/L \lesssim 0.1$.

The phenomena of slip and of temperature jump arise from the fact that molecules arriving at a stationary wall from a mean-free-path distant region of moving gas will carry with them the mean energy (temperature) and velocity characteristics of that region. When these are averaged with the properties of the molecules reflected from the wall, a finite-bulk gas slip velocity at the wall results, in contrast with the zero-slip velocity boundary condition employed with the Navier-Stokes equations in the continuum regime. Likewise, the temperature of the gas immediately adjacent to the wall is found to be different from that of the wall itself; thus the terminology temperature jump. The wall layer where these effects occur, the Knudsen layer, has been the subject of many detailed kinetic theory analyses based on the Boltzmann equation or models of it. These analyses have yielded Eqs. (32) and (33) for the fluid-solid

$$u_0 = C_m\lambda_0\left(\frac{\partial u}{\partial y}\right)_0 + \frac{C_s v_0}{T_0}\left(\frac{\partial T}{\partial x}\right)_0 \quad \text{(velocity slip)} \tag{32}$$

$$T_0 - T_w = C_t\lambda_0\left(\frac{\partial T}{\partial y}\right)_0 \quad \text{(temparature jump)} \tag{33}$$

surface boundary conditions at a plane wall ($y = 0$), with x and velocity u in the direction of the flow, where the subscript 0 indicates that the quantities are evaluated in the gas at $y = 0$. Currently accepted kinetic theory values for the coefficients are $C_m = 1.14$, $C_s = 1.1.7$, and $C_t = 2.18$, for diffuse reflection in a monatomic gas. The first term in Eq. (31) is the velocity slip due to finite wall velocity gradient, while the second is the thermal creep term arising from a temperature gradient in the gas in the direction of flow. Thermal creep is responsible for the phenomenon of thermophoresis, where small suspended particles experience a force in the direction opposite to a temperature gradient (although the phenomenon persists throughout the entire range of rarefaction, to the collisionless limit). Slip and temperature jump have the effects of reducing skin friction and heat transfer, but these effects are often obscured by other low Reynolds number and compressibility effects, and it has been found in general that slip flow corrections to the Navier-Stokes equation have a very limited range of applicability.

Transition flow. Numerous attempts have been made to develop higher-order continuum equations to replace the Navier-Stokes equations for regimes of moderate rarefaction. The results, which for the most part have been based on expansions of the Boltzmann equation, such as the Chapman-Enskog method, have proved not to be measurably superior to the Navier-Stokes equations. More success has been achieved through the use of simple interpolation between the slip flow and near-collisionless flow limits, or by solving the Boltzmann equation or models of it, or by direct simulation. This last method appears to be the most powerful and versatile of those mentioned, and will be described briefly.

The direct-simulation method models a real gas flow by computing the trajectories and positions of thousands of simulated molecules, and storing their positions and velocities as a function of time. With each time step, the molecules are advanced along their current trajectories, representative collisions with other molecules or with solid surfaces are calculated, the new velocities and positions are stored, and the process is repeated for the next time step. The flow is always unsteady, since at time $t = 0$ the molecules are introduced into the flow field in an arbitrary fashion, but a steady flow field is obtained as the long time average limit of the unsteady flow. The method has proved successful in modeling shock-wave structure, the free jet expansion of a gas into a vacuum, multidimensional flows about bodies, and polyatomic and chemically reacting gas flows. *See* SIMULATION. Lawrence Talbot

Bibliography. M. H. Aksel and O. C. Eralp, *Gas Dynamics*, 1993; J. D. Anderson, *Modern Compressible Flow: With Historical Perspective*, 3d ed., 2003; G. A. Bird, *Molecular Gas Dynamics and the Simulation of Gas Flows*, 1994; A. B. Cambel and B. H. Jennings, *Gas Dynamics*, 1958; S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases*, 3d ed., 1970, reprint 1990; J. E. A. John, *Gas Dynamics*, 2d ed., 1984; A. H. Shapiro, *The Dynamics and Thermodynamics of Compressible Flow*, vol. 1, 1953, vol. 2, 1954, reprint 1983; W. G. Vincenti and C. H. Kruger, *Introduction to Physical Gas Dynamics,* 1965, reprint 1975.

Gas field and gas well

The term gas field refers to a geographical area which is underlain by one or more commercial reservoirs of petroleum. This commercially valuable gas, primarily methane but with smaller amounts of ethane, propane, and butanes and in some cases containing significant concentrations of carbon dioxide, hydrogen sulfide, and nitrogen, is produced through wells which penetrate subterranean reservoirs, composed of porous rock. The gas in the reservoir may be free or dissolved in crude oil.

About one-third of the gross production of natural gas in the United States is produced from reservoirs in which there is no substantial amount of crude oil (mixture of higher-molecular-weight pentanes plus hydrocarbons) in contact with the gas. The gas in such reservoirs is called nonassociated gas. Associated gas is that produced from reservoirs in which there is a substantial contact of free gas (in the gas cap) with crude oil or gas dissolved in crude oil.

Nonassociated gas is produced as a result of volumetric expansion of the gas as the reservoir pressure is decreased from its original value down to the economically determined abandonment pressure. The recovery efficiency from such volumetric reservoirs is very high. The flow of associated gas through a reservoir and ultimately into the bore hole of a producing well is complicated by the fact that the two fluids are flowing simultaneously and essentially competing to reach the well bore. In fact, it is the release of gas from solution and the subsequent expansion of the free gas within the reservoir which provides the energy to drive the oil from solution gas reservoirs. It is usually necessary to restrict the production of natural gas from associated gas reservoirs in order to maximize oil production. *See* NATURAL GAS; OIL AND GAS FIELD EXPLOITATION; OIL AND GAS WELL DRILLING; PETROLEUM; PETROLEUM GEOLOGY; PETROLEUM RESERVOIR ENGINEERING. Todd M. Doscher

Bibliography. N. J. Hyne, *Nontechnical Guide to Petroleum Geology, Exploration, Drilling and Production*, 2d ed., 2001; W. C. Lyons, *Standard Handbook of Petroleum and Natural Gas Engineering*, 2d ed., 2004.

Gas furnace

An enclosure in which a gaseous fuel is burned. Domestic heating systems may have gas furnaces. Some industrial power plants are fired with gases that remain as a by-product of other plant processes. Utility power stations may use gas as an alternate fuel to oil or coal, depending on relative cost and availability. Some heating processes are carried out in gas-fired furnaces. *See* STEAM-GENERATING FURNACE.

Among the gaseous fuels are natural gas, producer gas from coal, blast furnace gas, and liquefied petroleum gases such as propane and butane. Crude industrial heating gases carry impurities that corrode or clog pipes and burners. Solid or liquid suspensoids are removed by cyclones or electrostatic precipitators; gaseous impurities are removed chemically. The cleaned gas may be mixed with air in the furnace, in the burners, or in a blower before going to the burners. The gas and air may be supplied at moderate pressure, or one or both at high pressure. The high-pressure component may serve to induce the other component into the furnace. The burner may be a single center-fire type or a multispud type with numerous small gas parts, depending on how the heat is to be concentrated or distributed in the furnace. Crude uncleaned gases are fired through burners with large ports, the burners being removable for cleaning. *See* FUEL GAS. Ralph M. Hardgrove

Gas thermometry

A method of measuring temperatures with gas as the thermometric fluid. Gas thermometry is the primary source of information about a fundamental physical parameter, temperature, over the range from about 3 to 900 K (−454 to 1160°F).

In principle, gas thermometry consists of using the ideal gas law, Eq. (1), where P is the pressure, V the

$$PV = nRT \tag{1}$$

volume, n the number of moles of gas, R the molar gas constant, and T the thermodynamic temperature, to evaluate an unknown temperature by reference to the single defining temperature of the Kelvin thermodynamic temperature scale, namely the triple-point temperature of water. This reference temperature, achievable within about one part in 10^7 by standard laboratory practice, has the value 273.16 K (0.01°C or 32.018°F). Determination of the unknown temperature requires two sets of measurements of the pressure P and the volume V of n moles of an ideal gas. One set of measurements usually is performed while the gas is maintained at the water triple-point temperature; the second set is obtained with the gas at the unknown temperature. The unknown temperature can readily be evaluated by rewriting Eq. (1) as $n = (PV)/(RT)$, which has the same value for both sets of measurements. The desired result is given by Eq. (2), where the primed quantities refer to the unknown temperature. In Eq. (2) the precise values of

$$T' = \frac{273.16(P'V')}{(PV)} \tag{2}$$

the gas constant R and of n need not be considered. *See* GAS; TRIPLE POINT.

Unfortunately, no real substance obeys the ideal gas law over any substantial range of temperature. Furthermore, the measurement of gas pressure and gas volume at the part-per-million level requires painstaking effort, as does the successful manipulation of measured quantities of high-purity gases.

Several different techniques have been developed to minimize the uncertainties arising from gas nonideality and from difficult measurements. These techniques include constant-volume gas thermometry, constant-pressure gas thermometry, constant-bulb-temperature gas thermometry, and isothermal gas thermometry.

Constant-volume gas thermometry. This is by far the most commonly used of the gas-thermometry methods. The name is somewhat misleading because no gas bulb truly exhibits constant volume over any substantial temperature range. Several steps are involved: inserting a fixed mass of a working gas into a rigid container or bulb; determining the pressure of the gas at the triple-point temperature of water; heating or cooling the container to a new temperature whose value is to be determined; and measuring the gas pressure at the new temperature.

Uncertainties. For best accuracy, many details in the construction and operation of the gas thermometer must be carefully pursued. These include accurate measurement of the working pressure; correction for the deviation from ideality of the working gas; minimization of impurities in the working gas, determination of the nature and concentration of the remaining impurities, and correction for their effects; determination of the change of volume of the gas bulb between the two operating temperatures; correction for the quantity of working gas that is not maintained at the working temperature (the so-called dead-space correction); correction for hydrostatic pressure effects; correction for the thermomolecular pressure effect; and evaluation of nonuniformity of the working temperature. Whenever mercury manometry is employed for gas-thermometer pressure measurements, as is often the case, additional uncertainties must be considered, particularly concerning the density of the mercury and the acceleration due to gravity at the measuring site. *See* MANOMETER.

In the determination of temperatures above the triple-point temperature of water, uncertainties arising from most sources are exaggerated approximately in the ratio of the Kelvin temperatures involved. The reverse is true for gas thermometry below room temperature.

The working gas used in most modern gas thermometry is helium, isotopically enriched to nearly pure ^{4}He. This gas is essentially inert chemically. Its departure from ideal thermodynamic behavior can be described by the virial equation (3). For work at

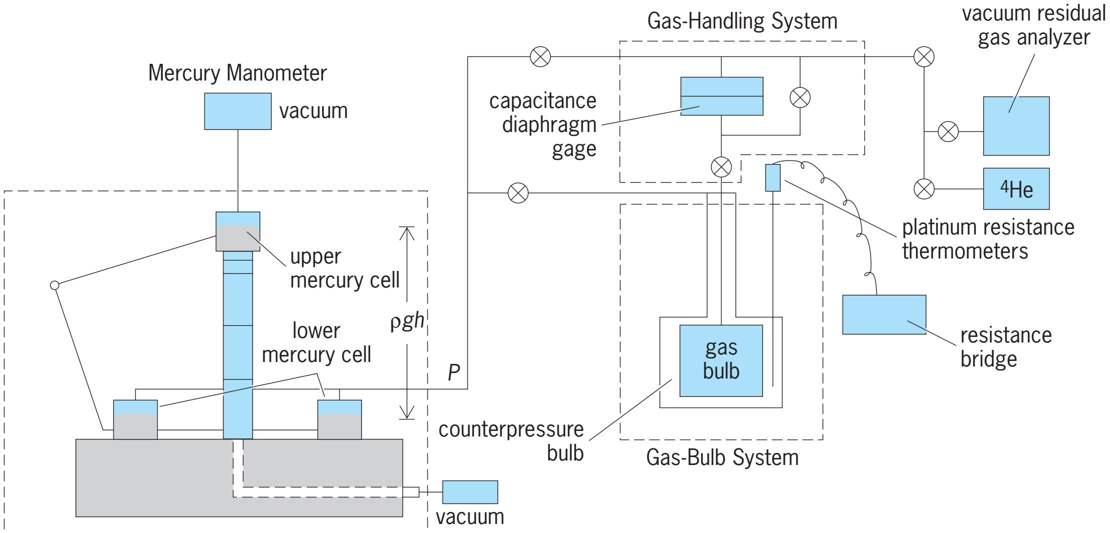
$$PV = nRT(1 + BP + CP^2 + \cdots) \tag{3}$$

the moderate pressures of 100 kilopascals (1 atmosphere) or less, usually used in gas thermometry, only B, the second virial coefficient, need be known. *See* VIRIAL EQUATION.

Temperature evaluation. The evaluation of unknown temperatures by constant-volume gas thermometry generally proceeds by equating the total number of moles of working gas at the two temperatures, as in Eq. (4). Here, the term with subscript b' accounts for the gas in the bulb at the unknown temperature. The term with subscript d' describes gas

$$\begin{aligned} n' &= \left[\frac{PV}{RT[1+BP]}\right]_{b'} + \left[\frac{PV}{RT[1+BP]}\right]_{d'} \\ &= n = \left[\frac{PV}{RT[1+BP]}\right]_{b} + \left[\frac{PV}{RT[1+BP]}\right]_{d} \end{aligned} \tag{4}$$

contained in the dead space during the unknown-temperature measurements; usually the dead-space gas is distributed over a range of temperatures between the laboratory temperature and the unknown temperature. The term with subscript b refers to gas contained in the gas bulb at 273.16 K. The term with subscript d refers to gas contained in the dead space during the triple-point measurements. Equation (4) indicates the inclusion of nonideality and dead-space effects; besides these, the effects of gas-bulb thermal



Typical constant-volume gas thermometer. (*After J. F. Schooley, Thermometry, CRC Press, 1986*)

expansion and the other problems mentioned above must be made part of the calculations.

Instrumentation. A typical constant-volume gas thermometer (see **illus.**) includes a mercury manometer, used for the measurement of pressure; a gas-bulb system; and a gas-handling system. The part-per-million accuracy achievable in manometric pressure measurements is central to the thermometric accuracy of this instrument. The entire manometer is operated in a temperature-controlled environment. The distance between the surface of the mercury in an upper cell and mercury surfaces in two lower cells is measured by the use of wrung stacks of calibrated end gage blocks. Axial holes through the gage blocks permit detection of the quality of the wringing process by measurement of the internal vacuum of the stack. The pressure exerted by the column of mercury is given by the product ρgh, where ρ is the density of the mercury, g is the acceleration due to gravity at the manometer, and h is the height of the gage-block stack.

The gas bulb is completely enclosed by a second bulb in which a so-called counterpressure of helium gas equal to the gas-bulb pressure is maintained at all times. The counterpressure gas minimizes pressure-induced changes in the gas-bulb volume and helps to reduce contamination of the working gas from impurities in the gas-bulb thermostat.

Isothermal gas thermometry. It is possible to determine the thermodynamic temperature of a gas bulb by repeatedly adding measured quantities of a working gas to it and measuring the resulting pressures. The repeated measurements at a single temperature are known as an isotherm. The virial equation (3) can be used to obtain both the unknown gas-bulb temperature and the value of the second virial coefficient.

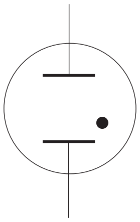
Two variations on the technique can be used, the absolute isotherm and the relative isotherm. In absolute isothermal gas thermometry, measured quantities of working gas are introduced into a gas bulb at the unknown temperature from a known volume that is maintained at 273.16 K. In the relative isotherm method, the working gas is added stepwise to a gas bulb while it is maintained at another, more convenient reference temperature. *See* LOW-TEMPERATURE THERMOMETRY; PHYSICAL MEASUREMENT; TEMPERATURE MEASUREMENT.

James F. Schooley

Bibliography. T. D. McGee, *Principles and Methods of Temperature Measurement*, 1988; T. J. Quinn, *Temperature*, 2d ed., 1991; J. F. Schooley, NBS/NIST gas thermometry from 0 to 660°, *J. Res. Nat. Inst. Stand. Tech.*, 95:255–290, 1990; J. F. Schooley (ed.), *Temperature, Its Measurement and Control in Science and Industry*, vol. 5, 1982; J. F. Schooley, *Thermometry*, 1986.

Gas tube

An electron tube, usually a diode (see **illus.**) or triode, containing a gas. The gas contributes two significant advantages: (1) The plasma, or gas of electrons and ions, that is formed in the gas during conduction allows space-charge effects to be overcome; and (2) the gas acts to reduce power dissipation at electrodes. These factors result in tubes that can conduct high currents and remain in use for long periods. Gas tubes are particularly important for high-power applications. Two important types are the thyratron and the gas-filled rectifier.



Symbol for a gas-filled diode.

Voltage and current characteristics. The voltage and current characteristics of a gas tube differ from those of a vacuum tube. In a vacuum tube, quasifree charge, or space charge, is present. The power dissipated in a vacuum tube (which is the product of the voltage drop times the current) is usually larger because of the electric fields produced by space charge. This is described by the Child-Langmuir law, which states that the current in a vacuum diode is proportional to the voltage to the $3/2$ power. In addition, the charges between electrodes acquire significant energy from the applied voltage and release this energy when they strike the electrode. These two means of energy loss are reduced in a gas tube for the following reasons. As electrical breakdown occurs in a gas tube, a plasma is formed. A plasma is a mixture of electrons and ions that may be electrically neutral overall, because the electrons are provided by the ions, and thus the net charge is balanced. When a potential is applied, the positive charges move in a direction opposite to that of the electrons. Thus, it is possible to overcome the obstacle of the large electric fields due to space charge, and provide more charges besides, by producing a plasma. Most current is carried by the electrons, because their mass is much less than even the lightest of positive ions, and they move much faster than a heavy ion in the same potential field. The gas also acts to slow down electrons because of collisions between electrons and un-ionized atoms or molecules within the plasma. This reduces energy loss at the electrodes. *See* ELECTRICAL CONDUCTION IN GASES; SPACE CHARGE; VACUUM TUBE.

Most of the potential fall in a tube occurs at the interface between the plasma and cathode. This region includes the cathode fall, which is a thin sheath that has a potential drop that corresponds to the energy required to ionize the gas in the tube. Thus, electrons are emitted from the cathode (requiring energy), and accelerated through a potential until they have sufficient energy to begin ionizing the gas that is present. This energy varies, depending on the ionization potential of the gas. *See* IONIZATION POTENTIAL.

Choice of gas. The gas is chosen with regard given to hold-off voltage and current-handling capabilities, as well as other factors related to specific applications, such as rate of rise of current (turn-on time), which is important for some pulsed applications. For example, mercury and xenon are used for high-current applications, and hydrogen is used for high-voltage applications.

Thermionic cathode. Many gas tubes use a thermionic or hot cathode, that is, a cathode that is heated and produces current through emission from a low-work-function material such as barium oxide. The development of improved cathodes remains an important technological problem. A good thermionic cathode allows a tube to operate for long periods with excellent repeatability and good current-handling capability. *See* THERMIONIC EMISSION.

Applications. Solid-state devices have replaced vacuum and gas tubes for many applications. Among the important remaining applications for gas tubes are high-power switching and pulsed-power applications requiring high power and rapid closing, or turn-on. For example, the hydrogen thyratron is still used as a radar modulator switch, and is also applied to the switching of high-power lasers. *See* LASER; RADAR.

Types of tubes. The hydrogen thyratron is a triode that operates at hydrogen pressures between 300 and 900 millitorr (40 and 120 pascals). A plasma is formed, and a glow discharge may be seen in a glass-enclosed tube during the closed or on phase. Hydrogen thyratrons are rated for stand-off voltages up to about 50 kV, and peak currents up to about 5000 A (off the shelf). They also close, or turn on, rapidly. They use a thermionic cathode and have a reservoir of hydrogen within the tube. They achieve a high stand-off voltage through an interesting “trick.” The grid is positioned very near the anode—about 0.1 in. (2 or 3 mm)—rather than away from it. The high stand-off voltage occurs because the average distance that an accelerated electron travels is comparable to the grid-anode separation, thus reducing the ionization of the gas, and therefore preventing breakdown. The stand-off voltage in one hydrogen thyratron may be 20 kV, but that tube can be switched on with only 100 to 200 V, applied between the grid and cathode. *See* BREAKDOWN POTENTIAL; ELECTRICAL BREAKDOWN; GLOW DISCHARGE.

Other gas-phase switches include the ignitron, where a pool of mercury serves as a cathode and mercury vapor is the conducting gas; the liquid-metal plasma tube, where a small amount of mercury is distributed over a metal cathode; the crossed-field tube, using a magnetic field to bend the paths of electrons; spark gaps; mercury rectifying tubes; and gas voltage regulators, which take advantage of a constant voltage across a glow discharge to clamp the voltage of an erratic source.

The spark gap has applications for very high voltages and currents. It operates in an arc mode, rather than a glow-discharge mode. The gas may be at atmospheric pressure, and gases such as nitrogen and helium are used. The spark gap tends to suffer from electrode degradation caused by arcing, but is able to stand off higher voltages and conduct higher currents than other gas-phase switches. *See* ARC DISCHARGE; SPARK GAP.

Outlook. Gas tubes are expected to retain their importance for higher-power applications, for example, voltages exceeding roughly 5 or 10 kV, along with high currents and fast turn-on. They are resilient to most perturbations. However, the continuing development of solid-state devices is expected to have increasing effect in this area as well: the thyristor (a *pnpn* semiconductor) has already made significant inroads. *See* SEMICONDUCTOR RECTIFIER.

Martin Gundersen

Bibliography. R. W. Crompton et al. (eds.), *Gaseous Electronics and Its Applications*, 1991; Institute of Electrical and Electronics Engineers, *IEEE Guide for the Application of Gas Tube and Air Gap Arrester Low-Voltage (Equal to or Less than 1000 Vrms* *or 1200 Vdc) Surge-Protective Devices*, 1993; J. A. Kunc and M. Gundersen, A fundamental theory for high-power thyratrons, I and II, *Lasers and Particle Beams*, vol. 1, pt. 4, pp. 395–425, 1983.

Gas turbine

One of a class of heat engines which use fuel energy to produce mechanical output power, either as torque through a rotating shaft (industrial gas turbines) or as jet power in the form of velocity through an exhaust nozzle (aircraft jet engines). The fuel energy is added to the working substance, which is gaseous in form and most often air, either by direct internal combustion or indirectly through a heat exchanger. The heated working substance, air comixed with combustion products in the usual case of internal combustion, acts on a continuously rotating turbine to produce power. The gas turbine is thus distinguished from heat engine types where the working substance produces mechanical power by acting intermittently on an enclosed piston, and from steam turbine engines where the working substance is water in liquid and vapor form. *See* INTERNAL COMBUSTION ENGINE; STEAM TURBINE.

All the variations of the gas turbine discussed below involve the transfer of heat to and from the gaseous working substance for the purpose of enhancing overall performance.

Air cycle. Gas turbine engines depend on the principle of the air cycle, just as their piston engine counterparts do, but they use a different succession of processes. In the basic simple cycle (**Fig. 1**), ambient air is first compressed to a maximum pressure level, at which point fuel heat energy is added to raise its temperature, also to a maximum level. The air is then expanded from high to low pressure through a turbine. The expansion process through the turbine extracts energy from the air, while the compression process requires energy input.

As the air moves through the engine, the turbine continuously provides energy sufficient to drive the compressor. In addition, because the turbine expansion process starts from a high temperature that comes from the fuel energy released by combustion, surplus energy beyond that required for compression can be extracted from the air by further expansion. At the point where the turbine has provided sufficient energy to power the compressor, the air pressure remains higher than the outside ambient level. This higher pressure represents available energy in the air that can be turned into useful output power by a final expansion process that returns the air pressure to ambient. The exhaust air leaves the engine with pressure equal to the outside, but at a higher temperature. As with any heat engine, the high exhaust temperature represents wasted energy that will dissipate into the outside atmosphere. *See* COMPRESSOR.

From an energy accounting standpoint, the sequence of processes acting on the air from front to rear constitutes a full cycle. It starts with the outside air entering at its initial state, and is completed when the air returns again to both ambient pressure and temperature levels. The series of cycle processes includes the final outside dissipation of the wasted exhaust energy, inevitable for every heat engine according to Carnot's principle. The ideal version of the gas turbine cycle is known as the Brayton cycle. *See* BRAYTON CYCLE; CARNOT CYCLE.

For any completed cycle, the total energy added from the fuel sources will always be equal to the sum of the useful output energy and the wasted exhaust energy. The thermal efficiency, which is the ratio of net output energy to fuel input energy for the cycle, measures the engine's ability to minimize wasted energy. A thermal efficiency of 60% means that for every 100 units of added energy 60 units will be available as useful output while 40 units will leave the engine as high-temperature exhaust.

Another performance measure is the specific power, which is the ratio of output power to quantity of working substance mass flow rate. Gas turbine engines, in comparison with other types of heat engines, are characterized not only by high levels of efficiency but also by very high levels of specific power. They are especially useful for applications that need compact power.

Simple cycle engines. By far the most common mechanical arrangement for the gas turbine is an in-line axial flow positioning of all components (**Fig. 2**). In the ground-based engine, the inlet at the front guides the incoming air into the compressor, which in turn delivers high-pressure air into the combustor section. The combustor burns the injected fuel at a high reaction temperature, using some of the air itself as an oxygen source. The combustion products in the combustor mix with the remaining unused air to reach a uniform equilibrium temperature, still high but diluted down from the reaction temperature. The hot, high-pressure combustor exit air enters the compressor drive turbine, where it expands down in pressure toward, but staying higher than, ambient level. This expansion process results in output shaft power that can be delivered directly

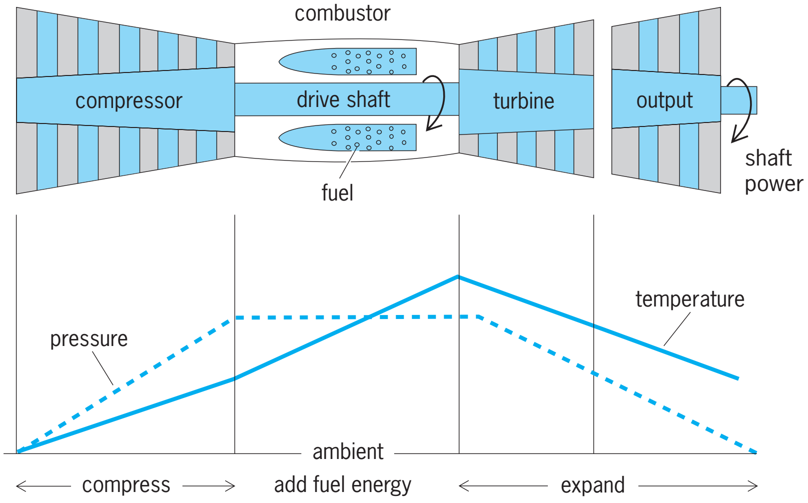


Fig. 1. Basic gas turbine cycle.

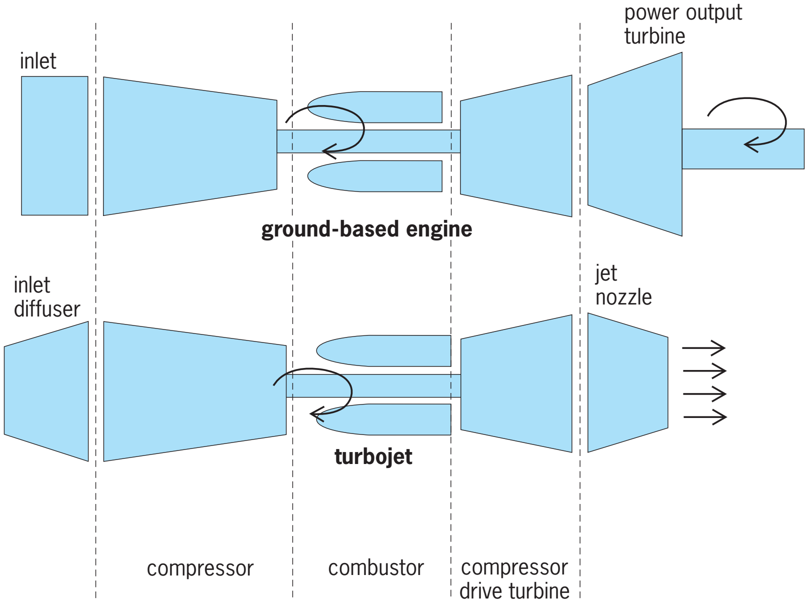
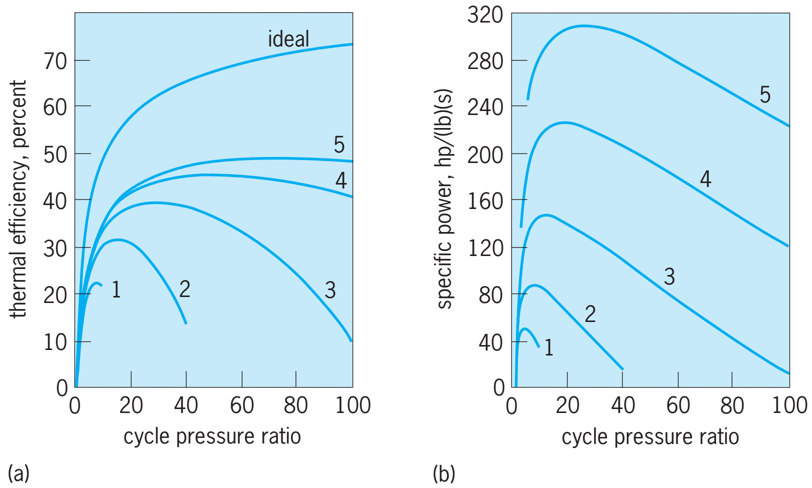


Fig. 2. Simple gas turbine component arrangements.

to the compressor through a connecting rotating shaft. Starting from the exit of the compressor drive turbine, net output power remains available. This power can be realized through the process of further pressure expansion completely down to the ambient level. For ground-based applications, the final expansion takes place through a power turbine whose output shaft is connected to the external load. In the single-spool arrangement the power turbine and compressor drive turbine are indistinguishably combined into one unit which, together with the compressor and the output load, is connected to a common shaft. For aircraft applications, either a power turbine extracts useful power to drive a propeller through a separate shaft (turboprop), or the expansion process takes place through a nozzle which acts to convert some of the thermal energy into velocity energy to be used for jet propulsion. *See* AIRCRAFT ENGINE; JET PROPULSION.



Turbine inlet temperature at standard day 59°F (15°C)

1 = 1000°F (537.8°C) 3 = 2000°F (1093.3°C) 5 = 3000°F (1648.9°C)

2 = 1500°F (815.5°C) 4 = 2500°F (1371.1°C)

Fig. 3. Simple cycle performance. (*a*) Thermal efficiency versus cycle pressure ratio and (*b*) specific power versus cycle pressure ratio, which is the ratio of maximum pressure to ambient pressure.

Design performance. For the cycle composed of ideal (100% efficient) components with no friction losses, thermal efficiency increases continuously with increased maximum pressure, expressed as a ratio to ambient pressure. Real components, however, temper this ideal trend. When cycle pressure ratio exceeds some value, the deleterious effects of component aerodynamic inefficiencies begin to predominate. This causes thermal efficiency to reach a maximum level at some pressure-ratio value and then fall off, depending on the levels of both the maximum cycle (turbine inlet) temperature and the component efficiencies. Increasing the level of maximum turbine temperature acts to offset the effect of energy waste due to component inefficiencies, as manifested by a shift in trend curves toward the ideal (**Fig. 3*a***). Specific power output increases directly with turbine entrance temperature and peaks with increasing pressure ratio, depending on temperature level (Fig. 3*b*).

Component elements. **Figure 4** shows a typical gas turbine engine in exploded view to reveal the individual component elements in a dual-spool arrangement.

Compressors. Axial-flow compressors are composed of a series of peripheral vanes designed to impart a swirling velocity to the incoming air as air density is increased. The individual vanes are shaped as airfoils to most efficiently turn the air away from the axial direction. They are arrayed in rows surrounding, and held in place by, rotating discs. The simultaneous turning and pressurizing of the air exerts a torque resistance that is overcome by an input shaft connected to the disc. Behind each row of rotating vanes, a row of stationary blades is emplaced that acts to straighten the swirling flow while increasing its pressure. This action of converting swirl velocity energy to air-pressure increase by means of a mechanical rotor enables continuously flowing air to be compressed. It is the central operating principle that makes feasible the gas turbine engine as a continuous power-producing machine.

Compressors are usually made up of several rows of alternating rotor vanes and stator blades, known as stages. By using multiple stages in tandem, the pressure of the air can be gradually increased by small increments, resulting in the most aerodynamically efficient process for achieving high overall pressure ratios. At the very highest of pressure ratios, the compressor is split into two or even three separate in-line compressors. Each is connected to its own shaft, nested one inside the other, and driven by individual turbines. The separate compressor and turbine set with its drive shaft is known as a spool. Placing two or more spools in series enables each group of compressor stages to operate near its own most aerodynamically efficient rotating speed. Progressing from front to rear, individual stages inherently

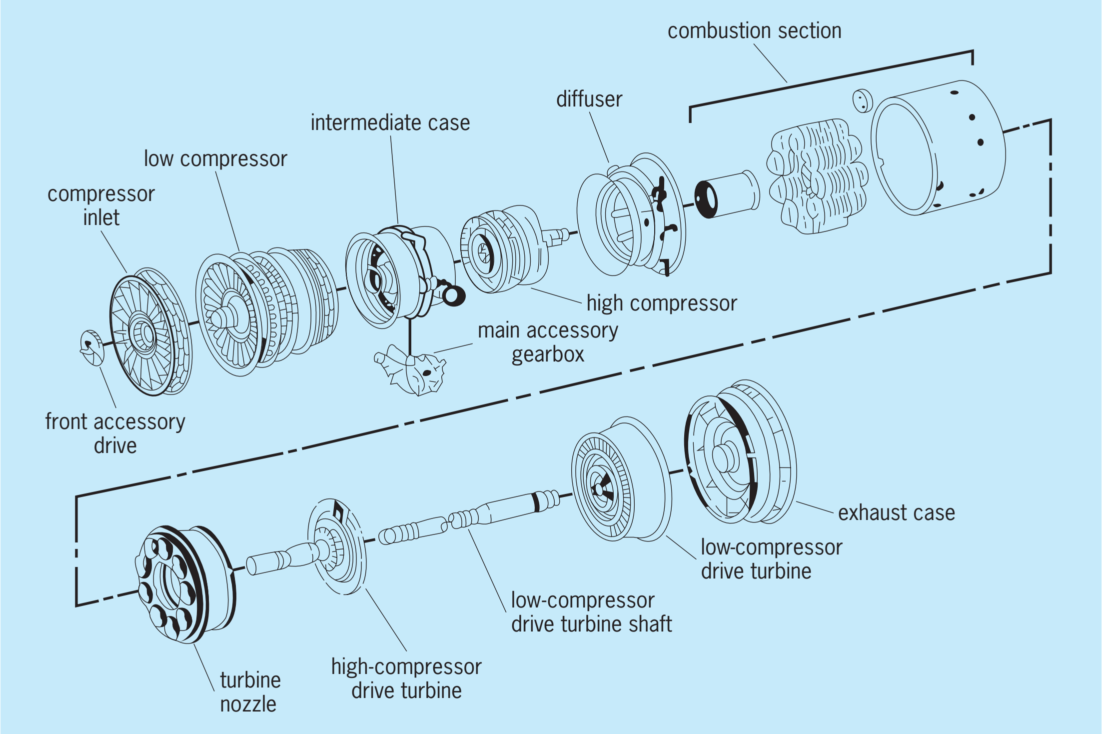


Fig. 4. Typical dual-spool gas turbine components.

function much better when free to travel at increasingly faster speeds.

Radial-flow compressors are most often found in very small gas turbine engines. In contrast to axial flow, the radial-flow, or centrifugal, compressor achieves air compression by imparting swirl to the air as it moves from the center of the rotor to its periphery. At the outer portion the swirling air energy is converted to increased pressure by means of fixed vanes that act to diffuse the air to lower velocity. Inherently, because the swirl process takes place while the air moves outward in radius, radial-flow compressors tend to present much greater frontal diameters for the same airflow quantity. As an offsetting advantage, they are simpler in design and less expensive to manufacture. This is especially important in the smaller sizes where radial-flow vanes are inherently less sensitive to those aerodynamic efficiency losses that must be controlled by stringent manufacturing tolerances in axial-flow airfoils.

A few gas turbines employ hybrid, or mixed-flow, compressors that, on a single spool, attach an axial-flow compressor directly to a radial-flow compressor.

Combustors. The gas turbine combustor is designed for simplicity and compactness while providing for efficient, stable combustion with low emissions. Clean combustion products always come with efficient burning. Combustion stability and efficiency are achieved primarily by significantly reducing the air velocity from the relatively fast level at which it leaves the compressor to a very slow level with high turbulence. An increasing-area diffusion duct placed between the two components acts to reduce air velocity. At the end of the diffuser, atomizing fuel nozzles spray a fuel mist into the low-velocity airstream where a high temperature flame front is established. To feed the combustion zone with oxygen, metal liners, pierced by a succession of small holes, direct the air inward toward the center of the combustion chamber where the spray nozzles are located. The rest of the air, entering at the cooler compressor exit temperature around the outside of the combustion zone, is forced radially inward by another series of small liner holes designed for thermal mixing. The process of mixing cold with hot continues downstream until the overall gas temperature distribution is sufficiently devoid of hot peaks to enter the turbine. *See* COMBUSTION CHAMBER.

Turbines. Axial-flow turbines are most commonly employed. Turbine operation is the reverse of that of compressors. Turbines are also composed of a series of stationary vane and rotating blade pairs, or stages, designed in the form of airfoils. However, the stationary turbine vanes act first to convert the thermal energy of the air into rotating swirl velocity as density decreases. The swirling air leaving the vanes impinges on the traveling airfoil blades and imparts a resisting torque to the vanes. The rotating airfoils change the swirl back to the axial direction while reducing air pressure in the process. Their torque travels through the discs holding the blades to the connecting output shaft that is attached to the resisting, power-absorbing load on its other end. In their most elementary version, turbines act just like simple pinwheels.

Radial in-flow turbines are usually coupled with radial-flow compressors. This accommodates the large diameters inherent to both in a back-to-back arrangement. Radial-flow turbines reverse the operating principle of their compressor counterparts. Surrounding the periphery, high-temperature air is expanded to high velocity in a series of nozzles that directs swirling air inward toward the rotating vanes. As the air moves toward the center, the vanes expand the air to lower pressure while its swirl decreases. The air with reducing swirl exerts a tangential force on the vanes that appears as torque acting on the connecting shaft.

Critical to the progression of gas turbine technology, turbine airfoils must withstand increasingly high temperatures along with strong centrifugal and bending stresses. They are made from metal alloys with protective coatings, both especially developed for aircraft engines, and they employ interior cooling. Complex internal passages circulate cooler air drawn from the compressor upstream through the inside of the turbine airfoil. Small, laser-drilled holes lead the inside air through blade surfaces into the outside main stream. *See* TURBINE.

Diffusers/nozzles. Air that travels through ducting responds in velocity and pressure to changes in area along the duct. If the area gets smaller, pressure goes down while velocity speeds up; if the area gets bigger, pressure increases while velocity slows down. Area-increase conduits, called diffusers, are integral gas turbine components. A diffuser is used to slow down the compressor air entering the combustor. In aircraft gas turbines a diffuser is also placed in front of the compressor to slow down the captured outside air from forward-flight velocity to a velocity that is aerodynamically compatible with the front compressor airfoils. The consequent pressure increase in the inlet diffuser, or ram pressure, contributes to the overall cycle pressure ratio. This means that to achieve the same maximum pressure ratio, flight gas turbines powering high-speed airplanes do not need as high a pressure ratio from the compressor as do low-speed ones. However, supersonic aircraft must be equipped with complex inlet diffusers to achieve efficient ram-pressure increases.

Area-reduction conduits are called nozzles. Nozzles in flight gas turbines find use in the exhaust section. The pressure level exiting the compressor drive turbine expands through the engine exhaust nozzle down to ambient pressure while the jet velocity increases to provide jet propulsion.

Heat exchangers. The components described above are found in all gas turbines from the simple to the very complex. The more complex cycles use heat exchangers either to remove heat to cool the inside air, using an outside cooling source, or to add heat from a hotter outside source. In gas turbines the external cooling sink usually is a nearby body of water such as a lake or a river. In the water-air case a shell-and-tube arrangement is most practical and economical. In the case of recuperators or regenerators (see below), the heat exchange normally takes place between two separate streams of air. In this instance the preferred heat exchanger would be the plate-fin type. As a rule, cycles that employ heat exchangers to improve output and efficiency are limited to only those special power plant applications that warrant their use on the basis of overall economics. Air, even in a compressed state, has such relatively low density that the size and cost of the heat exchanger alone can outweigh any performance advantages. *See* HEAT EXCHANGER.

Gas turbine variations. A number of gas turbine engines have been developed that use variations of the simple cycle system.

Closed versus open systems. A gas turbine's working substance can either pass through the engine continuously with its exiting exhaust products replaced by the incoming air (open), or it can recycle the air (closed). Because there will always be wasted exhaust energy, causing the gas to leave the turbine at the rear with a higher temperature than that of the air entering the compressor at the front, closed cycles require a cooling heat exchanger to be placed between those two components. This heat exchanger serves to dissipate the wasted exhaust heat into the outside atmosphere. It normally uses circulating water as the coolant and a cooling tower to dispose of the water's heat to the outdoors. The gas-cooled nuclear reactor system represents a prime example of a closed system. Here the working substance is helium gas. Helium passing through the reactor is impervious to radioactivity. An advantage of all closed systems is their ability to pressurize the fully contained gas. The high density reduces the size of all components.

Regenerative cycle. With a recuperating heat exchanger placed behind the turbine, some of the wasted exhaust energy can be recovered. This is accomplished by directing the compressor exit air through one side of a heat recovery unit while passing the turbine exhaust through the other side. The compressed air is externally ducted to the rear of the engine, receives the exhaust heat from the turbine side in a heat exchanger, then is returned to reenter at the front of the combustor. The higher-temperature air entering the combustor requires less fuel to reach the same maximum cycle temperature and output power (**Fig. 5**).

Intercooled cycle. Advantage can be taken of a basic physical principle governing the air-compression process; that is, the power needed to compress is directly proportional to the entrance temperature. When a cooling heat exchanger is placed in the middle of the compressor to lower air temperature at that point, the net work needed to power the back part of the compression process is accordingly reduced. This net reduction in compressor work becomes directly available as useful power that can be added to the output turbine (Fig. 5).

Intercooled-regenerative cycle. The principles that govern regenerative and intercooled cycles can be combined into a single cycle. Here, the two separate principles reinforce each other in such a way that the overall thermal efficiency becomes higher than for either of the two separate cycles (Fig. 5).

Reheat cycle. Strictly for the purpose of increasing output specific power, adding a second combustor

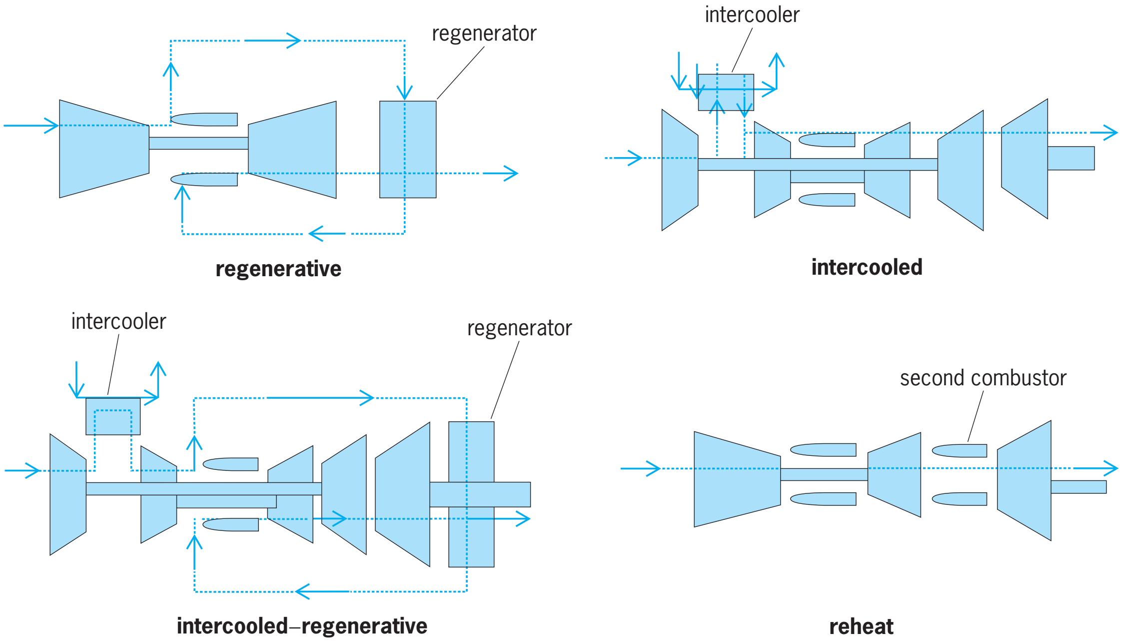


Fig. 5. Gas turbine variations.

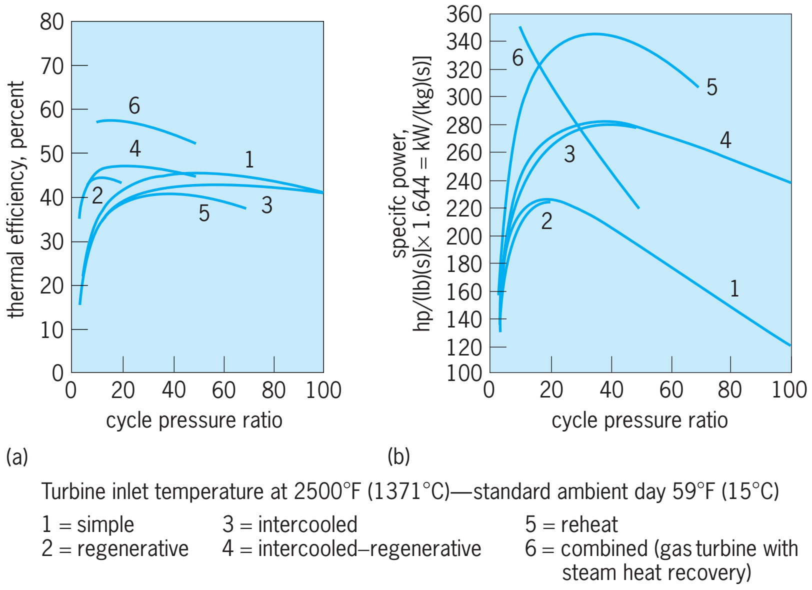
behind the turbine allows for extra fuel energy input within the limitations of the turbine to withstand excessive temperature. When the heat addition point is placed behind the compressor drive turbine, the power output turbine can be conveniently located behind the second combustor as a single unit. This arrangement avoids having an internal drive shaft penetrate the hot combustor section from either end. In aircraft applications the second combustor is known as an afterburner. The afterburner is particularly valuable for jet fighters that need short, sudden bursts of thrust power (Fig. 5).

Externally fired system. Although not common, in certain circumstances processes other than internal combustion can be the source of a gas turbine's energy addition (firing). The nature of the fuel source most often dictates the need for external combustion, with a heat exchanger to transfer heat indirectly to the air. The nuclear reaction, for which the air must be shielded from radioactivity, is one example. Coal, unacceptable for internal combustion in a gas turbine due to its ash residue, is another.

Design performance of simple cycle variants. For all of the variants as well as the basic simple cycle, both thermal efficiency and specific power are related strictly to the maximum turbine-inlet temperature (related to local ambient temperature as a ratio) and the maximum pressure (related to local ambient pressure as a ratio). **Figure 6** compares typical efficiency and specific power trends as a function of these key variables. The trends depicted are also related to the level of efficiency and pressure losses and turbine cooling air penalties associated with each component. Figure 6 assumes typical levels.

Industrial applications. The industrial gas turbine industry started in the late 1950s, following the development of aircraft gas turbines. The three applications of industrial gas turbines are generation of electric power, mechanical drive (principally gas or oil pipeline compression), and marine propulsion. Over 90% of gas turbine applications, as measured in megawatts (MW) of power produced, are in electric power generation. The majority of the navies of the world use gas turbines to propel most of their surface ships.

In electric power generation service, gas turbines offer the combined advantages of lower capital cost, shorter installation time, high efficiency, and low environmental impact as compared to steam



Turbine inlet temperature at 2500°F (1371°C)—standard ambient day 59°F (15°C)

1 = simple
2 = regenerative
3 = intercooled
4 = intercooled–regenerative
5 = reheat
6 = combined (gas turbine with steam heat recovery)

Fig. 6. Typical performance for gas turbine variants, where cycle pressure ratio is the ratio of maximum pressure to ambient pressure.

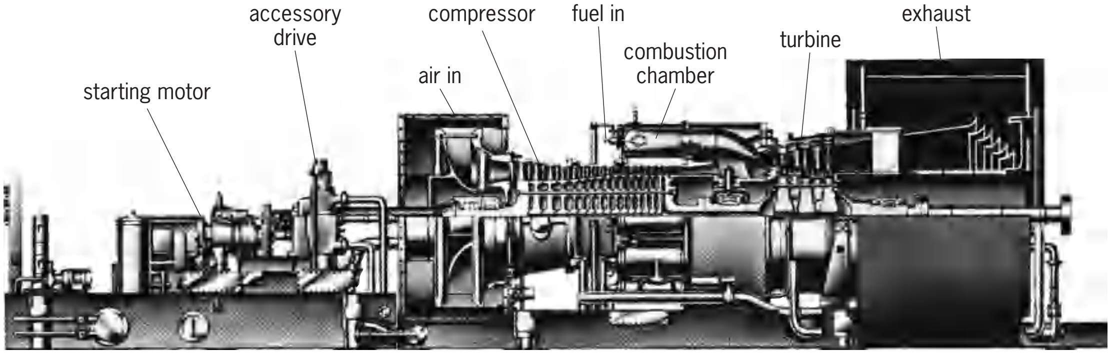


Fig. 7. Industrial gas turbine.

turbine–based power plants. As a result, gas turbine–based power plants have captured a majority of the market for new power generation equipment. The U.S. Energy Information Agency, in its Annual Energy Outlook for 1998, projected that 85% of new power generation equipment to be added worldwide by the year 2020 would be gas turbine–based.

Compliance with emissions regulations, particularly that for nitric oxide, is a major application consideration. In the late 1970s and the 1980s, water or steam was injected into the combustion reaction zone to decrease the flame temperature, which is the principal parameter affecting nitric oxide emission. Present technology is producing combustion system designs that feature premixed air and fuel in lean mixtures to reduce the flame temperature without any outside diluent.

Industrial design. Most gas turbines with outputs above 1–2 MW have multistage axial-flow compressors and turbines. Lower-power units tend to have single-stage centrifugal compressors and radial-inflow turbines to minimize weight, size, and cost. Thermal efficiency improves with larger size, as friction and tip leakages become a smaller percentage of power produced, and multistage axial-flow components become more efficient than radial stages. **Figure 7** shows typical design features for a large frame-type gas turbine designed specifically for industrial application.

The same major components are evident in **Fig. 8**, which shows an aeroderivative gas turbine, that is, one which was derived from an aircraft jet engine. The engine depicted has a twin-spool (dual-compressor) arrangement. Hot gases exiting the gas generator portion drive the power turbine, which is connected to the external load. Typical aeroderivatives differ from frame-type gas turbines by being lighter in weight and having higher design pressure ratios. Their outputs are limited to a maximum of about 50 MW because of the size of the aircraft jet engines from which they are derived. The largest frame-type gas turbines have outputs of over 200 MW. Aeroderivatives tend to produce higher simple-cycle efficiencies, whereas frame types produce the highest combined-cycle efficiencies. For a given output, frame-type gas turbines tend to be somewhat lower in cost than the aeroderivatives.

For applications where the output shaft speed varies, as in a compressor drive, a separate power turbine is needed. This requires a multiple-shaft gas turbine such as shown in Fig. 8. A single-shaft gas turbine has a typical effective operating range of 85–105% of rated speed. With a separate power turbine to drive the load, the output shaft speed range is typically about 50–105% of rated speed.

Typical performance. Component efficiencies, airflow, pressure ratio, and turbine inlet (or firing) temperature are the major factors affecting gas turbine output and efficiency. Typical multistage axial component efficiencies are in the 86–93% range. Material developments and turbine cooling techniques permit turbine rotor inlet temperatures to exceed 2500°F (1370°C) for the most advanced units. The **table** shows typical thermal efficiencies and pressure ratios for various cycles representing engines manufactured and in actual service.

Aircraft engines. Invented simultaneously in England and Germany in the late 1930s expressly for airplane power, the aircraft gas turbine as a jet engine has grown to become the dominant power plant for both

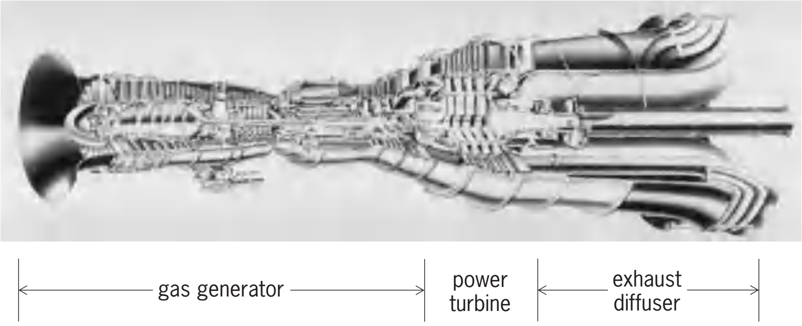


Fig. 8. Aerodervative gas turbine. (*Pratt & Whitney*)

Typical pressure ratios and thermal efficiencies

Cycle	Pressure ratio	Thermal efficiency, %
Simple	15–30	32–42
Intercooled	40–60	45–49
Regenerative	8–10	37–40
Combined cycle	15–23	50–60

commercial and military aircraft. Its use for ground-based applications started in the 1950s when, with their recognized compact power and quick start-up capability, these engines began to be adapted directly from aircraft engine designs. The use of jet propulsion to replace piston-driven propeller planes allowed airplanes to take an incremental step in flight speed. Jet engines offered two significant advantages. First, the gas turbine's power output for a given size far surpasses that of the piston engine. Second, and most important for flight, jet propulsion overcomes the propeller's most severe limitation. As flight speeds approach the sonic velocity, airplane propellers drastically lose aerodynamic efficiency. Because of their consequent superior speed capability, jet fighters, as they evolved from the early experimental versions in the latter part of World War II, soon replaced the propeller-driven planes with their inherent speed limitation. Subsequently, this became true for commercial passenger flight as well. Even use of the compact gas turbine to power propellers in what is called the turboprop engine is no longer widespread. With the further development of supersonic flight, the jet engine became the only choice as the power plant for these aircraft. *See* TURBOJET.

Jet propulsion is achieved when exhaust gases are ejected at high velocity. As an opposite reaction in force to the engine's pushing the air backward, the air pushes the engine forward. The gas turbine develops the high exit velocity as a result of expansion of the gases in the exhaust nozzle. Leaving the compressor drive turbine at a pressure level higher than ambient, the exiting gases further expand to ambient pressure in the rear nozzle by a process that converts the latent available exhaust thermal energy directly to energy of motion, that is, kinetic energy.

As an evolution of the original simple turbojet, most jet engines are of the bypass turbofan variety. Turbofans more efficiently convert available power to propulsive power. In a sense, the turbofan is a hybrid combination of pure turbojet and propeller engines. After extraction of turbine power sufficient to drive the compressor, the remaining useful energy of the working air is divided in two. First, some additional energy is extracted by further expansion in an additional turbine. This added turbine drives a fan that, like a propeller, adds velocity to air that surrounds and bypasses the engine. By propelling the bypass air rearward, the fan provides a reactive forward thrust force. The remainder of the available energy then further expands in the jet nozzle to produce thrust power. The major difference between the fan and the propeller lies in the external ducting which surrounds the bypass air. The turbofan's ducting contains the inlet air as it is slowed by diffusion to a much lower velocity than the flight velocity. For the air entering the bypass fan as well as the front compressor, this internal diffusion avoids the efficiency degradation that the propeller suffers with high aircraft speeds. *See* TURBOFAN.

Finally, small gas turbines with power turbines connected directly to the lifting rotor are frequently used as power plants for helicopters.

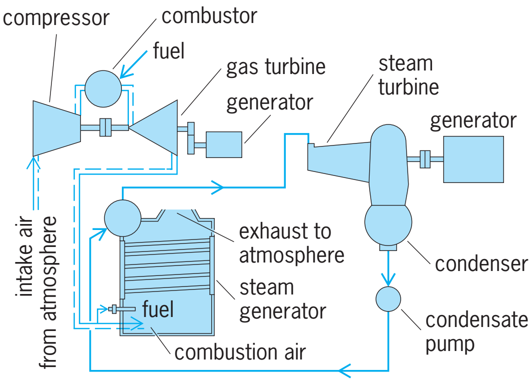


Fig. 9. Combined cycle.

Automotive use. Some of the major automobile manufacturers have experimented with gas turbines for cars and trucks. However, any advantages over the ubiquitous piston engine have not proven sufficient to displace the worldwide infrastructure of garages, gas stations, etc. already in place to support the existing automobile piston engine. However, the U.S. Army's Main Battle Tank (M1A1, Abrams) is powered by a small gas turbine engine. To accommodate the need for variable shaft torque at widely different traveling speeds and climb gradients, these vehicles require a gearing system between the engine's output shaft and the ground traction system. Such would be the case for any road vehicle.

Hybrid power-plant systems. In addition to variations of the simple-cycle gas turbine described in the previous section, hybrids of gas turbines and other power-plant equipment are used in the electric power generation industry. The most common of these is the combined cycle, in which the exhaust gas from the gas turbine is passed through a boiler to make steam for a steam turbine (**Fig. 9**). In some cases, the boiler has supplemental heat from a separate fuel source, as indicated in the figure.

Another industrial variant of the simple cycle is cogeneration. The term refers to the use of the exhaust heat from the gas turbine for some process other than electric power generation. The most common arrangement for cogeneration is to make steam from the exhaust heat, as in a combined cycle, but the steam is sent to the process instead of to a steam turbine. Examples of processes to accept the steam can be heating or domestic hot-water systems for buildings or industrial processes in chemical plants. *See* COGENERATION SYSTEMS.

Much work has been done to adapt gas turbines to coal fuel. Coal gasification, utilized to remove contaminants, is particularly attractive since the gas turbine combined cycle can be integrated into the gasification process for improved efficiency. Air extracted from the gas turbine can be used as the source of oxidant for the gasification process, and the steam produced in the process can be expanded through a steam turbine. **Figure 10** shows how various components relate in an integrated gasification combined

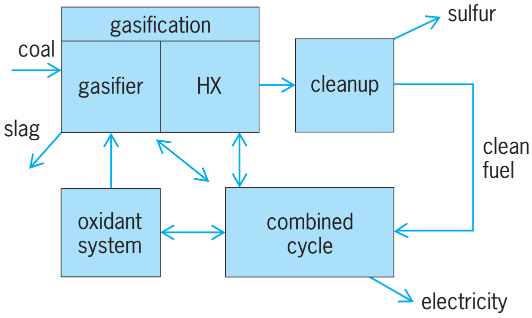


Fig. 10. Integrated coal gasification system.

cycle (IGCC). Petroleum coke is also used as feedstock for an IGCC. In some cases, biomass (for example, wood chips) has been used as fuel for an IGCC.

Industrial fuels. Aviation gas turbines use special jet fuels that come directly from petroleum refineries. Ground-based gas turbines widely use natural gas. It is both clean-burning and easy to transport, if a supply pipeline is nearby. Lower-cost liquid fuels, typically number 2 home heating oil or diesel oil, can be used in industrial applications. Most engines employed for electric power generation are equipped with dual fuel capability, operating on either natural gas or liquid fuel as needed. Such is the case in regions where, during periods of cold weather, home heating takes highest priority for the available natural gas supply, or where fluctuating local natural gas prices cause the use of liquid fuel to be variably more or less economical.

The other two applications for industrial gas turbines use one fuel each: gas from the pipeline for pipeline compressor drives, and liquid fuel for marine propulsion. Some use of various solid fuels has been accomplished in gas turbines, as described above in the section covering hybrid types.

Controls and part-load operating characteristics. Gas turbines characteristically produce smooth and linear throttle response over their entire operating range. Rotor speeds normally vary continuously over this range without the need for the gear shifting and clutch mechanisms found in piston engines. The governing fuel control senses rotor speeds, pressures, and temperatures to maintain stable, steady power or thrust output and, when needed, ensure rapid accelerations and decelerations. The control is programmed, normally by electronic input, to guard against harming the engine during throttle changes by governing the appropriate fuel input rate. Most important, during throttle transients the control functions to prevent turbine overheating, burner blowout, and compressor surge.

The gas turbine output shaft speed can be flexibly adapted to the speed characteristics of a wide variety of types of external loads. While many engines mate with their external loads through a separate shaft, or free power turbine, the flexibility feature holds true even for single-shaft designs. The free turbine arrangement can most efficiently adapt to the widest range of load characteristics. However, in the extreme, single-shaft industrial engines used for electric power generation can directly connect to and drive alternating-current electric generators that must maintain constant speed from idle to full power.

Part-load efficiency normally falls off from the design condition in varying amounts. This fall-off occurs least in the regenerative cycle variant which, for applications such as marine propulsion and automobiles, can offer a compelling advantage. Both ships and ground transportation vehicles, where regeneration is common, tend to consume the most fuel when cruising for long periods of time well below their maximum accelerating power.

John H. Lewis; William H. Day

Gasket

Deformable material used to make a pressure-tight joint between stationary parts, such as cylinder head and cylinder, that may require occasional separation. Gaskets are known as static seals, as compared with packing or dynamic seals. In packings the parts are frequently in motion, as in piston rods and valve stems. *See* PRESSURE SEAL.

Gaskets are made of sheet materials such as natural or synthetic rubber, cork, vegetable fiber such as paper, asbestos and plastic pastes, or of soft metallic materials such as lead and copper. Rubber in the form of O-rings is used for light pressure.

Paul H. Black

Bibliography. P. H. Black and O. E. Adams, Jr., *Machine Design*, 3d ed., 1968; R. C. Juvinall, *Fundamentals of Machine Component Design*, 3d ed., 2000; J. E. Shigley and C. R. Mischke, *Mechanical Engineering Design*, 2d ed., 1989; Society of Automotive Engineers, *SAE Handbook*, revised annually.

Gasoline

A mixture of hydrocarbons whose boiling point is below 200°C (390°F), obtained in the fractional distillation of petroleum. Gasoline is a liquid at ambient temperature, but it volatilizes readily in air to form a flammable mixture. The hydrocarbon fuel is used to power the internal combustion engine. Gasoline is composed primarily of the alkanes (paraffins) hexane, heptane, and octane, plus smaller amounts of higher-boiling alkanes. *See* ALKANE; INTERNAL COMBUSTION ENGINE.

Historical development. The major force behind the existence of the petroleum industry is the constant consumer demand for products such as gasoline and other liquid fuels. There have been many changes in emphasis on product demand since petroleum first came into use some 5000–6000 years ago. Most of the changes relate to the establishment of the modern refining industry prompted by the discovery of petroleum in the United States in 1859. The main

Component streams for gasoline

		Boiling range	
Stream	Producing process	°C	°F
Paraffinic			
Butane	Distillation Conversion	0	32
Isopentane	Distillation Conversion Isomerization	27	81
Alkylate	Alkylation	40–150	105–300
Isomerate	Isomerization	40–70	105–160
Straight-run naphtha	Distillation	30–100	85–212
Hydrocrackate	Hydrocracking	40–200	105–390
Olefinic			
Catalytic naphtha	Catalytic cracking	40–200	105–390
Steam-cracked naphtha	Steam cracking	40–200	105–390
Polymer	Polymerization	60–200	140–390
Aromatic			
Catalytic reformate	Catalytic reforming	40–200	105–390

demand at that time was for the kerosine fraction, used as an illuminant. Thereafter, the development of the internal combustion engine brought about a demand for a new product. *See* KEROSINE.

Production. In the early days of the industry, gasoline requirements were met by distilling crude oil, that is, by maintaining the oil at high temperatures until the different oil components (fractions) reach boiling point and vaporize. The fraction obtained was referred to as straight-run gasoline. However, the heavier (more viscous) crude oils do not contain any gasoline, and the equivalent fraction (naphtha) is now produced by thermal decomposition (cracking), which involves decomposition of the higher-molecular-weight constituents under temperatures greater than 450°C (840°F).

Gasoline is usually produced by catalytic cracking or by reforming processes. In catalytic cracking, the petroleum (or petroleum-derived feedstock) is fed into a reaction vessel containing a catalyst. In reforming, naphtha (refined or unrefined) is heated with hydrogen in the presence of a catalyst. Reforming causes a rearrangement of the structures of the molecular constituents and creates a gasoline product. *See* CRACKING; PETROLEUM PROCESSING AND REFINING; REFORMING PROCESSES.

Composition. The hydrocarbon constituents in the boiling range of gasoline are those that have 4–12 carbon atoms in their molecular structure. Thus, gasoline can vary widely in composition; even gasolines with the same octane number may be quite different. For example, low-boiling distillates with high (above 20%) aromatics contents can be obtained from some crude oils. The variation in aromatics content as well as the variation in the content of normal paraffins, branched paraffins, cyclopentanes, and cyclohexanes is dependent upon the characteristics of the petroleum feedstock, and influence the octane number of the gasoline. *See* AROMATIC HYDROCARBON; DISTILLATION.

Into the first decade of the twentieth century, the gasoline produced was that which was originally present in crude oil or which could be condensed from natural gas (so-called natural gasoline). However, it was soon discovered that if the heavier portions of petroleum (such as the fraction that boils higher than kerosine, for example, gas oil) was heated to more severe temperatures, thermal degradation (or cracking) occurred to produce smaller molecules that were within the range suitable for gasoline. Therefore, gasoline that was not originally in the crude petroleum could be manufactured. *See* NATURAL GAS.

Blends. The differences in composition of gasoline dictate that, in order to produce a uniform product, blending of the products from several component streams (see **table**) is necessary. The properties of each stream may vary considerably, significantly affecting the product gasoline. The blending process is relatively straightforward, but the determination of the amount of each component to include in a blend is much more difficult. The operation is carried out by simultaneously pumping all the components of a gasoline blend into a pipeline that leads to the gasoline storage. The pumps adjust for the correct proportion of each component, while baffles in the pipeline are often used to mix components as they flow to the storage tank.

The low octane number of many paraffinic naphtha fractions can limit their use as gasoline components, although other properties may make them desirable. Other complications arise through the compositional changes induced in the component streams by processing variations. For example, an increase in cracking temperature will produce a higher-octane cracked naphtha (often due to the presence of aromatic constituents), and before this naphtha can be included in a blend, adjustments must be made in the proportions of the other hydrocarbon components.

Volatility. Volatility is an important property of gasoline and is a necessity to ensure engine starting in cold weather. In winter, volatility is raised and the flash point is lowered by adding the more volatile butanes and pentanes. To prevent vapor lock in warm weather, the amounts of the more volatile constituents are reduced to produce mixtures that will not vaporize in the fuel lines.

Aviation gasoline, used in light aircraft as well as certain types of civil aircraft, has a narrower boiling range [38–170°C (100–340°F)] than automobile gasoline [0–200°C (30–390°F)]. The narrower boiling range ensures better distribution of the vaporized fuel through the more complicated induction systems of aircraft engines. Since aircraft operate at altitudes where the prevailing pressure is less than the pressure at the surface of the Earth, the vapor pressure of aviation gasolines must be limited to reduce preliminary vaporization (boiling) in the tanks, fuel lines, and carburetors. *See* AIRCRAFT ENGINE; AIRCRAFT FUEL.

Additives. Additives are incorporated into commercial gasoline blends, for example to inhibit oxidation and gum formation during storage. Dyes may be added for identification purposes. Alcohol and surfactants are used to reduce carburetor icing and corrosion. Detergent additives remove from the engine and fuel injector some of the deposits produced by gasoline combustion.

Octane number. Before they are ignited by a spark plug, the hydrocarbons in a gasoline blend may ignite spontaneously under the high temperature and pressure conditions inside an engine cylinder. This preignition causes a characteristic engine knock.

The octane number is a measure of the ability of a hydrocarbon fuel to resist preignition. It is obtained by comparing the antiknock performance of the gasoline with that of a mixture of isooctane and heptane: a gasoline blend with an octane number of 90 equals in performance a mixture of 90% octane and 10% heptane. The octane number of a gasoline can be increased by the use of reforming techniques and by alkylation, where gasoline components are recombined to build a larger molecule with a high octane number. *See* OCTANE NUMBER.

Gasohol. Gasohol is a mixture of gasoline and ethyl alcohol that is used as a motor fuel. Although ethyl alcohol is a petroleum product, the necessary commercial quantities for gasohol production can be obtained by fermentation of agriculture waste. Methyl alcohol is also suitable for gasohol mixtures.

Standard automobiles can use gasohol containing up to about 17% ethyl alcohol, although 10% is usually recommended because of environmental concerns. To use higher percentages, an automobile engine must have modifications made to its carburetor, spark plugs, engine timing, and seals. Gasohol burns more slowly, coolly, and cleanly than gasoline, providing greater octane and fewer pollutants. A liability of gasohol is the greater solvent properties of ethyl alcohol, which in concentrated amounts can damage rubber seals and diaphragms and certain finishes. *See* ALCOHOL; ALCOHOL FUEL.

Environmental aspects. For many years tetraethyllead was added to gasoline to increase octane ratings and improve fuel efficiency. It was found, however, that the extremely high levels of airborne lead in urban areas were directly traceable to the combustion of leaded gasoline. In 1985 the U.S. Environmental Protection Agency ruled that gasoline lead content be reduced by over 90%, and by the early 1990s unleaded gas had become the standard throughout the United States. As a result, there was a dramatic drop in airborne lead emissions. To control the levels of other polluting emissions from automobile exhausts, engines must be equipped with various emission-control devices. The most effective is the catalytic converter, which can operate only with unleaded gasoline. *See* AIR POLLUTION; CATALYTIC CONVERTER; PETROLEUM. James G. Speight

Bibliography. J. G. Speight, *Gas Processing: Environmental Aspects and Methods*, 3d ed., 1999.

Gasteromycetes

An artificial class of fungi in the phylum Basidiomycota in which basidiospores are produced in a mass (gleba) and enclosed within a membrane called the peridium. Such enclosed fruit bodies evolved many times; hence the class is heterogeneous. Most Gasteromycetes have lost the ability to discharge spores ballistically off basidia directly into air. Since internal sporulation prevents normal spore drop into air currents, these fungi have developed other dispersal mechanisms. Puffballs (Lycoperdales) typically form powdery spores in a perforated peridium that acts like a bellows when hit by rain. Earthstars (*Geastrum*) disperse similarly, but the peridia are elevated on arched rays. *Bovista* scatters spores while tumbling in the wind; *Calvatia* cracks open in place. In the gleba, cottony threads called capillitia slow the rate of dispersal. Stalked puffballs (Tulostomatales and Podaxales) characteristically form underground, emerging on stalks after spore maturation; they are adapted to desert life by their protected hypogeous maturation. False puffballs (Sclerodermatales) typically have thicker peridial walls than true puffballs.

The highly evolved Nidulariales have funnel-shaped structures that hold seedlike, splash-dispersed, sporulating disks called peridioles. These fruit bodies resemble miniature bird nests. The related *Sphaerobolus* violently everts an inner membrane, tossing a sticky glebal ball several centimeters. Stinkhorns (Phallales) mature their sticky, foul-smelling glebas in a hypogeous egg from which erupt showy phallic, floral or cagelike receptacles designed to attract insects. Flies or beetles ingest the spores and disperse them. Hymenogastrales (and possibly Melanogastrales and Gauteriales) are dispersed like truffles (Ascomycetes) as spores in feces after excavation and ingestion by animals lured to these false truffles by characteristic odors. *See* ASCOMYCOTA.

Gasteromycetes are either saprophytic or mycorrhizal. Historically, mature dry spores of puffballs were used to clot wounds. Young puffballs are edible, but false puffballs are poisonous. Because of their phallic shape, stinkhorns have been associated with witchcraft. *See* BASIDIOMYCOTA; EUMYCOTA; FUNGI.

Scott A. Redhead

Bibliography. M. B. Ellis and P. J. Ellis, *Fungi Without Gills (Hymenomycetes and Gasteromycetes): An Identification Handbook*, 1990; D. S. Hibbett et al., *Proc. Natl. Acad. Sci.*, 94:12002-12006, 1997; O. K. Miller and H. H. Miller, *Gasteromycetes: Morphological and Developmental Features*, 1988.

Gasterosteiformes

An order of teleost fishes composed of about 257 species, most of which have a small mouth at the end of a long tubular snout and body armor of dermal plates or bony rings. They are further identified by the absence of supramaxillary, orbitosphenoid and basisphenoid bones; and a pelvic girdle that is not connected to the cleithra. Gasterosteiforms are carnivorous predators, feeding chiefly on crustaceans, which they siphon from the water with their tubelike mouthparts. Large species such as trumpetfishes and cornetfishes are primarily piscivorous. Sticklebacks, pipefishes, and seahorses are popular aquarium fishes.

Gasterosteoidei (Thoracostei). The members (three families and about 10 species) of this suborder are identified by a protractile upper jaw, presence of circumorbital, lachrymal, nasal, and parietal bones, absence of the postcleithrum, and kidneys that produce a gluelike substance that is used by males to construct a nest from plant material.

Hypoptychidae (sand eel). A monotypic family that is known from marine waters of Japan and Korea north to the Sea of Okhotsk.

Aulorhynchidae (tubesnouts). Two coastal marine species comprise this family, one on the coasts of Japan and Korea, the other from Alaska to California.

Gasterosteidae (sticklebacks). There are five genera and no less than seven species of sticklebacks, the exact number awaiting studies of the various species complexes. Sticklebacks, so called because of a row of 3 to 16 free (not attached to each other by membranes), stiff, sharp spines that precede the soft dorsal fin. Additional identifying characteristics are a single spine in the anal and pelvic fins (the latter fin fails to develop in several species in certain localities); the mouth is not at the end of a tubular snout and the small jaws are equipped with teeth; a narrow caudal peduncle; and elaborate courtship behavior, which has been the subject of numerous studies. The body of marine species is armed with heavy bony plates, whereas the body of freshwater species is virtually naked. The maximum length is about 18 cm (7 in.). Sticklebacks inhabit marine, brackish, and freshwaters of the north temperate zones of the Northern Hemisphere. The male constructs a nest and provides parental care. *See* STICKLEBACKS.

Syngnathoidei (Solenichthyes). The fishes of this suborder have a small mouth at the end of a long tubular snout and the upper jaw is not protractible. Additionally, the pelvic fins when present are abdominal; the circumorbital bones, other than the lachrymal, are absent; ribs are absent; and the anterior three to six vertebrae are elongate. There are eight families and about 247 species in the suborder.

The following three families comprising the infraorder Syngnatha have uniquely shaped gill filaments. The filaments are reduced to small rosettelike tufts attached to rudimentary arches. This is called the lophobranch pattern (from the Greek *lopho*, "tuft," and *branchos*, "gill", and the fishes are referred to as the Lophomorphs.

Pegasidae (seamoths). This is a family of tropical marine shorefishes (up to 150 m (495 ft) in depth) of the Indo-West Pacific. The body is broad and depressed and encased in body plates; the mouth is beneath a long flattened rostrum; the pectoral fins are large, horizontal and supported by unbranched rays. There are two genera and five species, the largest of which attain 14 cm (possibly 18 cm) in length (5.5, possibly 7.1 in.).

Solenostomidae (ghost pipefishes). This is a family of only one genus and two species, each of which is known from tropical marine waters of the Indo-West Pacific. The family is distinguished by a short, compressed body with two separate dorsal fins, the first five weak spines, the second of small and posteriorly placed opposite the anal fin. The female has a brood pouch formed by the large pelvic fins. Unlike male pipefishes and seahorses, female ghost pipefishes care for the eggs. Maximum size is 16 cm (6.3 in.).

Syngnathidae (pipefishes and seahorses). The body of syngnathids in this family is elongate and encased in bony rings; the gill openings are small; the supracleithrum is absent; a kidney is present only on the right side; the pelvic fins are absent; dorsal, pectoral, and anal fins are usually present, the last is very small, and one or more of the three fins may be absent in adults of some species; the caudal fin is absent is some species and when absent the tail is prehensile. The male cares for the eggs, which are attached by the female to a brood organ (often a pouch) on the underside of his trunk or tail. Most species occur in the temperate and tropical shore waters of the Atlantic, Indian, and Pacific oceans; however, some New World species range into cool waters from southwestern Alaska to Tierra del Fuego. Syngnathids are primarily marine and brackish-water species; however, some species are known only from freshwater.

Two subfamilies are named, Syngnathinae (pipefishes) and Hippocampinae (seahorses); however, some genera are morphological intermediates, perhaps evolutionary links, between pipefishes and seahorses. Examples of intermediate forms are the seadragons of Australia, which are large and have ornate leaflike appendages. The body is not straight as in typical pipefishes, neither is it contorted to the extent seen in seahorses. Some of the intermediate forms are called pipehorses (see **illustration**).

There are about 190 species of pipefishes in

Alligator pipefish, *Syngnathoides biaculeatus*. (*Copyright © J. E. Randall*)

51 genera and only 25 species of seahorses, all in the genus *Hippocampus*. Seahorses differ from pipefishes by having the head bent 90° downward, a prehensile tail, no caudal fin, and they are limited to a marine environment. The pipehorses, in having a caudal fin, are placed with the pipefishes.

Indostomidae. This family, in the infraorder Indostomoida, is represented by a single species, *Indostomus paradoxus*, from southeastern Asia, where it inhabits swamps with soft bottoms and dense vegetation. It has a slender body covered with bony plates; the gill filaments are lobe shaped as opposed to comblike or tufted; the dorsal and anal fins each have six rays and are mirror images; the dorsal fin is preceded by five isolated spines, the opercle is armed with six spines; and the caudal peduncle is very long and slender.

Infraorder Aulostomoida. The following four families, comprising the infraorder Aulostomoida, have normal comblike gill filaments and the postcleithrum is present. Species of the families Aulostomidae and Fistulariidae share the following characteristics: The four anteriormost vertebrae are elongate; well-developed bones (nuchal plates) are medially placed at the nape; and the pelvic fins usually have six soft rays. Species of the families Macroramphosidae and Centriscidae have five or six elongate anterior vertebrae, and pelvic fins with one spine and four soft rays.

Aulostomidae (trumpetfishes). In this family trumpetfishes are tropical marine fishes, most commonly found on reefs of the Atlantic and Indo-Pacific oceans. They have an elongate, scaly, and compressed body (flattened laterally), a fleshy barbel at tip of the lower jaw, rounded caudal fin, and 8 to 12 free spines preceding the dorsal fin, which is a mirror image of the anal fin. These fishes lie at odd angles, most often in a head-down posture, as they wait to attach prey from ambush. A length of 80 cm is the maximum of the three species, all in the genus *Aulostomus*.

Fistulariidae (cornetfishes). Cornetfishes are also tropical marine fishes of reefs and shallow shore waters of the Atlantic and Indo-Pacific oceans. They have an elongate and depressed body (flattened dorsoventrally), the tubular snout is exceptionally long, the skin is either naked or with minute prickles and a linear series of scutes, no barbel at tip of the lower jaw, forked caudal fin with trailing elongate filament, no dorsal spines, and the dorsal fin a mirror image of the anal fin. Cornetfishes are piscivorous, feeding in open water as well as in coral reefs. The largest of the four species in the one genus, *Fistularia*, attains a length of 1.8 m; the remaining three species are usually less than 1 m.

Macroramphosidae (snipefishes). This family, composed of three genera and about 12 species found in the tropical and subtropical Atlantic, Indian, and Pacific oceans, is characterized by a deep and compressed body, usually with bony plates on the back, and the second of four to eight dorsal spines very long. Additionally the long and slender snout ends in a very small mouth, the jaws of which lack a barbel. The snipefishes are bathydemersal (on the bottom in deep water) marine occupants of the continental shelf and slopes to 1000 m (3300 ft), where they feed on benthic invertebrates.

Centriscidae (shrimpfishes). The shrimpfish family, composed of two genera and four species, are the most bizarre fishes among the gasterosteiforms, if not all fishes of the world. They are extremely compressed with the body almost entirely encased by thin bony plates that are expansions of the vertebral column. The elongate body, so encased and with an exceptionally sharp ventral keel, suggests a razor. The spikelike first dorsal spine is at the very end of the body, and it is inline with the flat dorsal contour of the body, displacing two short spines and closely spaced soft dorsal and caudal fins to a ventral position at the posterior end of the fish. Shrimpfishes, limited to marine and brackish waters of the Indo-Pacific, swim in a head-down vertical position, in schools or small groups among long-spine sea urchins and branching coral, to depths of 20 m to 100 m, depending on the species. They feed primarily on planktonic crustaceans and reach a maximum length of 15 cm (6 in.). Herbert Boschung

Bibliography. M. A. Bell and S. A. Foster (eds.), *The Evolutionary Biology of the Threespine Stickleback*, Oxford University Press, 1994; C. E. Dawson, Family Syngnathidae, in *Fishes of the Western North Atlantic*, Mem. Sears Found. Mar. Res. Mem. 1(pt. 8), pp. 1–172, 1982; P. C. Heemstra, Macrorhamphosidae, pp. 459–461, in M. M. Smith and P. C. Heemstra (eds.), *Smith's Sea Fishes*, Springer-Verlag, Berlin, 1986; E. S. Herald, From pipefish to seahorse—a study of phylogenetic relationships, *Proc. Calif. Acad. Sci.*, Ser. 4, vol. 29, pp. 465–473, 1959; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, 1994; T. W. Pietsch, Evolutionary relationships of the sea moths (Teleostei: Pegasidae) with a classification of gasterosteiform families, *Copeia*, 1978(3):517–529, 1978; R. J. Wootton, *The Biology of the Sticklebacks*, Academic Press, 1976; R. J. Wootton, *A Functional Biology of Sticklebacks*, University of California Press, 1984.

Gastornithiformes

An order of extinct flightless birds known from Paleocene and Eocene deposits in North America, Europe, and Asia. They are believed to have dispersed across the North Atlantic via a land bridge when those continents were connected in the Eocene. These were giant, 2-m-tall (7-ft) birds with large heads, huge laterally compressed bills, broad cervical vertebrae, wide pelves, massive legs with relatively short tarsi and heavy toes suggestive of a slow-moving gait, and reduced wings that were evidently too short to permit flight.

Mode of life. The mode of life of gastornithiforms is a matter of controversy. They have often been depicted as predators that evolved to fill the bipedal carnivore niche vacated by theropod dinosaurs at the end of the Cretaceous Period. A biomechanical analysis of their powerful jaw apparatus, suggesting considerable biting strength, has been used to support the idea that they were carnivorous and capable of feeding on both live prey and carrion. According to an alternative view, their jaw and hindlimb anatomy indicates that they were primarily herbivorous and incapable of sustained rapid running. Various skull and postcranial characters suggest that gastornithiforms are related to the waterfowl (Anseriformes) and, more distantly, to the true fowl (Galliformes). *See* ANSERIFORMES; GALLIFORMES.

Fossil record and classification. Gastornithiformes were first described from the Eocene of France in 1855, with the genus *Gastornis*. In 1876, a comparable giant bird from the Eocene of New Mexico was described as *Diatryma*. The original material from both Europe and North America was fragmentary, and *Gastornis* was first reconstructed in a highly inadequate fashion. When much more complete material from North America was discovered in the early twentieth century, it seemed that the North American form, *Diatryma*, was significantly different from the European *Gastornis*. Later, new specimens from Europe were described as *Diatryma*. A reexamination of the original French specimens and descriptions of new finds from Europe show that *Gastornis* and *Diatryma* were in fact very similar and should be placed in a single genus, *Gastornis* (the earlier name having priority).

Gastornis is known from the Paleocene of Germany, France, and Belgium; the Lower Eocene of France, England, the United States, and Canada; and the Middle Eocene of Germany. Two species (*Gastornis giganteus*, *Diatryma regens*) are recognized from the New World, and there may be three from the Old World (*G. parisiensis*, *G. russelli*, and *G. sarasini*). A second genus, *Zhongyuanus*, was described from the Lower Eocene of China on the basis of scanty material and remains poorly known. *See* EOCENE; PALEOCENE.

Habitat. Gastornithiforms are most abundant in alluvial sediments of the Bighorn Basin of Wyoming—where the most complete known skeleton, the holotype of *Diatryma steini* Matthew and Granger [=*Gastornis giganteus* (Cope); see **illustration**], was discovered; in fluvial deposits in eastern France; and in lake and swamp deposits of the Geiseltal of Germany. Associated sediments, floras, and faunas indicate that the Gastornithiformes inhabited well-vegetated, temperate-to-subtropical coastal lowlands and inland floodplains, much like the marshy habitats now occupied by their vegetarian relatives, the Anhimidae or screamers (Anseriformes), in South America.

Gastornis giganteus **(reconstruction). (*American Museum of Natural History*)**

Resemblance to modern birds. Modern avifaunas contain no exact counterparts to the gastornithiforms. According to the herbivorous interpretation, the closest analogs probably are the takahe, *Porphyrio mantelli* (Rallidae), and kakapo, *Strigops habroptilus* (Psittacidae) of New Zealand, which are flightless and herbivorous and are, respectively, the largest rail and parrot. Large body size and flightlessness in gastornithiforms, *Porphyrio*, and *Strigops* seem related to adoption of a low-energy diet of leaves, which in turn requires a large alimentary canal, forfeiture of the energetic expense of flight, and the evolution of a jaw apparatus specialized for feeding on vegetation. Supporters of the carnivorous hypothesis stress resemblances with some modern scavenging birds, such as the marabou stork (*Leptoptilus crumeniferus*) or the African carrion crow (*Corvus crassirostris*); however, these birds are not flightless. According to this interpretation, possible extinct analogs may be the South American phorusrhacids, which were giant flightless birds for which a predatory lifestyle is generally accepted. *See* AVES.

Allison V. Andors; Eric Buffetaut

Bibliography. A. V. Andors, *Diatryma* among the dinosaurs, *Nat. Hist.*, 104(6):68–71, 1995; A. V. Andors, Paleobiology and relationships of the giant groundbird *Diatryma* (Aves: Gastornithiformes), *Acta XX Congressus Internationalis Ornithologici*, vol. I, pp. 563–571, 1991; A. V. Andors, Reappraisal of the Eocene groundbird *Diatryma* (Aves: Anserimorphae), *Science Series*, Los Angeles, no. 36, pp. 109–125, 1992; L. D. Martin, The status of the Late Paleocene birds *Gastornis* and *Remiornis*, *Science Series*, Los Angeles, no. 36, pp. 97–108, 1992.

Gastrointestinal tract disorders

Malfunctions of the organs of digestion from the esophagus to the rectum. The gastrointestinal tract usually functions unnoticed, intruding into consciousness only in the form of specific sensations of hunger, intestinal motion, or the need to defecate. However, it is subject to a wide array of pathologic states.

Esophagus. The esophagus is the site of two principal disorders: disturbance of the muscular contractions, and mechanical interference with the passage of food caused by a narrowed lumen. Carcinoma of the esophagus produces mechanical obstructions to swallowing; the long-term prognosis is poor. Excessive acid reflux or regurgitation from the stomach may result in inflammation of the lower esophagus with heartburn and indigestion. Changes in diet and eating habits may bring relief.

Achalasia. The prime example of muscular or neuromuscular disorder is achalasia, a condition of unknown etiology characterized by degeneration of the nervous mechanisms that coordinate contractions of the esophagus and relaxation of the lower esophageal sphincter. Swallowing solids and liquids is difficult, and nutritional consequences usually result. The condition is normally painless, but some persons have active contractions of the body of the esophagus, a condition known as vigorous achalasia. Treatment is directed at reducing the pressure in the lower esophageal sphincter to permit food to pass by force of gravity. Weakening of the sphincter is accomplished surgically by cutting the outer circular muscle layer of the esophagus or by forcible pneumatic dilation with a balloon which ruptures the circular muscular fibers. Both techniques are reasonably effective. Medications such as calcium channel blockers improve motility and reduce pressure in the lower esophageal sphincter.

Connective tissue disease. Other disorders that affect the esophagus include connective tissue diseases such as scleroderma or dermatomyositis. Scleroderma is characterized by weakening of the smooth musculature, which is present in the lower two-thirds of the esophagus, and the lower esophageal sphincter. Muscle contraction is weakened, and the sphincter fails to prevent the reflux of gastric acid; reflux esophagitis frequently ensues. Dermatomyositis affects the striated musculature present in the upper third of the esophagus. There are no specific treatments for connective tissue diseases affecting the esophagus. However, in the case of scleroderma, acid reduction and antireflux measures are often helpful; surgery can reduce the hazard of acid reflux. *See* CONNECTIVE TISSUE DISEASE.

Mechanical disorders. Mechanical problems include stricture or scarring of the lower esophagus, which usually develops from acid reflux. A stricture forms slowly, usually preceded by symptoms of esophagitis, that is, heartburn and discomfort in the lower chest, particularly at night or in recumbency. Stricture tends to be progressive, with swallowing becoming increasingly difficult. It is diagnosed with a barium meal x-ray or by direct visualization with upper gastrointestinal endoscopy. Strictures can be dilated with cylindrical instruments called bougies, which are introduced orally in progressively larger diameters. Results are good with acid reduction and antireflux measures. Medications such as bethanechol and metoclopramide improve the tone of the lower sphincter and aid esophageal emptying. In severe cases, surgery may be required to reduce acid reflux.

Swallowing is also impeded mechanically by a lower esophageal, or Schatzki, ring, which is a mucosal constriction at the junction of the esophagus and stomach in the presence of a hiatus hernia. Reflux esophagitis may favor the development of such a ring, which tends to stabilize; the resulting problem is not progressive but intermittent and dependent upon the size of the food bolus. It may produce the steak house syndrome, in which a person who is eating, and possibly under the influence of alcohol, swallows a lump of food that sticks at the Schatzki ring. Treatment is based primarily on patient education, with occasional dilation or surgical correction of the tight ring. *See* HERNIA.

Cancer. Carcinoma of the esophagus causes mechanical obstructions to swallowing. The cancerous growth, whether arising in the squamous epithelium of the esophagus or in the columnar epithelium of the upper stomach, may progressively reduce the size of the lumen and lead to increasingly severe dysphagia. Ultimately, even swallowing fluids becomes impossible. Carcinoma of the esophagus or the upper region of the stomach is usually diagnosed by means of fiber-optic endoscopy and biopsy. The long-term prognosis for all forms of esophageal cancer is poor, but encouraging results have been achieved in Japan with high-dose radiation followed by extensive surgical resection. Palliative treatments are available: radiation therapy can temporarily shrink squamous cell cancer of the esophagus.

Reflux esophagitis. Reflux esophagitis is an inflammation of the lower esophagus that results from excessive acid reflux or regurgitation from the stomach. The common symptoms of heartburn and indigestion are produced by reflux of acid. The major cause seems to be a weakness in the lower sphincter, which ordinarily serves as a physiological barrier to stomach acid. Pressure within the sphincter may be low in the presence of a hiatus hernia; however, reflux esophagitis can occur even in the absence of hiatus hernia or other disease. Scleroderma may impair the lower esophageal sphincter and lead to reflux esophagitis. Other factors that may contribute include inadequate clearing of the acid from the esophagus, and the composition of gastric contents that reach the esophagus. Symptoms vary from occasional heartburn to severe distress caused by regurgitation of gastric acid into the throat, known as water brash, or even expectoration of acidic fluid. Treatment is directed at neutralizing gastric acid or reducing its secretion in the stomach, or at improving the tone of the lower esophageal sphincter. The individual can reduce acid reflux by elevating the upper body during sleep, eating smaller meals, retiring on an empty stomach, losing weight, and avoiding caffeine, tobacco, and foods found to cause distress. Those measures are usually successful; surgery is rarely required unless serious complications occur, such as bleeding from a highly inflamed esophageal lining or stricture and scarring.

Stomach. The stomach is the site of secretion of the highly concentrated hydrochloric acid. The lining of the stomach and duodenal both need to be protected against the effects of the acid.

Peptic ulcer disease. Duodenal ulcer is one of the most common maladies, but for unknown reasons its incidence seems to be declining. It is associated with high concentrations of hydrochloric acid caused by excessive gastrin secretion. The role of genetic or environmental influences is not known. In persons with a gastric ulcer, acid concentrations are usually normal or low. Gastric ulcer occurs in persons of middle age or older, whereas duodenal ulcer tends to be a disease of young adults. The cause of gastric ulcer is not known but may involve increased reflux of bile and other constituents of the small intestine. Bile salts are capable of breaking the gastric mucosal barrier, which normally keeps hydrochloric acid in the lumen. When the barrier is broken, hydrogen ions diffuse into the cells and cause inflammation. The barrier can also be broken by aspirin and other nonsteroidal anti-inflammatory drugs. The combination of aspirin and alcohol seems to be particularly deleterious: alcohol causes vascular congestion, and aspirin breaks the mucosal barrier. Together they increase the risk of bleeding in the upper gastrointestinal tract.

Peptic ulcer disease is typically diagnosed with an upper gastrointestinal barium meal or fiber-optic endoscopy. When an infectious agent, particularly *Campylobacter pylorii*, is suspected, it may be identified by biopsy of the lining of the stomach antrum. However, it is a difficult organism to culture, and methods of eradication are still being studied. Duodenal ulcer tends to be recurrent and even chronic; once the diagnosis is confirmed, further testing is usually not warranted. The presence of a gastric ulcer should be firmly established, since it can occasionally mask cancer. Highly effective medications are available for the treatment of duodenal ulcer. Antacids are useful in large doses. Histamine receptor blockers inhibit the secretion of hydrochloric acid from the parietal cell. Some drugs act at the site of the ulcer to provide a protective barrier against pepsin. Pharmaceuticals that inhibit the proton pump, which takes part in the final stage of parietal cell functioning, would dramatically reduce gastric acid secretion.

Surgery, once frequently performed for peptic ulcer disease, is now uncommon, and removal of part of the stomach is indeed rare. Surgery, when required, usually involves sectioning (cutting) of the vagal nerves or their branches to reduce gastric acid secretion. Such an operation is known as a vagotomy or selective vagotomy and is usually performed when noninvasive medical treatment has failed.

Gastritis. Inflammation of the stomach lining, known as gastritis, ranges in severity from acute to chronic and can ultimately lead to atrophy of the tissues and loss of function. Chronic forms of gastritis may occur in association with other diseases such as diabetes mellitus or hypothyroidism. When gastric atrophy is severe, gastric acid and pepsin are no longer secreted. In the most severe forms of gastritis, intrinsic factor is no longer secreted, and pernicious anemia with vitamin B_{12} deficiency develops. Immunologic factors seem to play an important role in pernicious anemia; antibodies to both parietal cells and intrinsic factor are detectable in the blood. *See* ANEMIA; DIABETES; THYROID GLAND DISORDERS.

Tumors. The incidence of gastric cancer has been declining more rapidly than any other cancer. At one time it was among the most common causes of death but is now less common than either colon or pancreatic cancer. Some of the decline has been attributed to improved refrigeration methods and decreased exposure to nitrites, which are thought to be carcinogenic. Nitrites are found in partially decomposed meats, smoked and cured meats, and cigarette smoke. Nonetheless, cancer of the stomach remains a serious problem. In Japan, where the incidence is high, screening programs have detected cancers in very early stages, when cure rates are high. The typical symptoms of gastric cancer include early satiety after eating small amounts of food, abdominal pain, anemia from blood loss, and difficulty in swallowing. Benign stomach tumors are usually asymptomatic but may cause bleeding.

Small intestine. The small intestine is normally resistant to disease. When nutrition is adequate, structure and function remain intact into old age with only slight changes if no specific diseases occur.

Malabsorption syndrome. A prototypical disease of malabsorption is gluten enteropathy, or nontropical sprue. Probably due in part to genetic abnormalities, gluten enteropathy is characterized by inflammation and loss of the normal architecture of the small intestine following ingestion of some proteins. In the extreme form, villus atrophy, the fingerlike projections from the intestinal wall are absent. The causative substance is contained in gluten, the protein of cereal grains; a gluten fraction, gliadin, contains the toxin that is responsible. A gluten-free diet excludes wheat, barley, rye, and oats, although the toxicity of oat protein has been questioned. Several genetic markers are evident in patients with gluten enteropathy, and antibodies to wheat protein fractions can be found in their blood. Diagnosis of the condition can be based on the response to gluten restriction and on a small-intestine biopsy performed with an endoscope or a swallowed capsule. A gluten-free diet, once it has been adopted, is usually maintained for life. In children the condition is called celiac disease; the diet must be observed throughout life despite lessening of the symptoms with advancing age.

Other conditions may also interfere with the absorptive function of the small intestine. Tropical sprue, found in tropical and subtropical climates, resembles gluten enteropathy in its manifestations. It may, however, be an infectious diarrhea; it can be treated with antibiotics and high doses of folic acid. A gluten-free diet does not cure the condition.

Infections within the small intestine, particularly those of the parasite *Giardia lamblia*, may also cause malabsorption. Bacterial infection by *Escherichia coli* causes most cases of traveler's diarrhea. Among the many other diseases that may lead to malabsorption is regional enteritis, or Crohn's disease. *See* INFLAMMATORY BOWEL DISEASE.

Tumors. Of all the body's organs, the small intestine is one of the most resistant to neoplastic disease. Benign tumors are rare, as are cancer and lymphoma. (In some familial conditions, however, benign polyps may develop.) An understanding of that resistance might have value in cancer research.

Colon. The colon, or large intestine, is subject to malfunction of widely varying degrees, ranging from mild irritation to life-threatening diseases.

Altered bowel function: constipation and diarrhea. Disturbance of bowel function is among the most common complaints. Constipation probably predominates, particularly with advancing age, but loose stool and diarrhea are also frequent and troublesome. A familiar condition associated with altered function is irritable bowel syndrome, a nonspecific label for an intestinal disturbance without known anatomic cause.

Constipation may be defined as bowel movements less frequent than once every 3 days or as hard stool that is difficult to pass. Until recently, the typical Western diet contained low amounts of fiber compared with that common in underdeveloped nations and probably contributed to the incidence of constipation. Fiber in the diet, particularly from the indigestible component of grains known as bran, frequently gives relief. In some cases, a severe problem seems to worsen with age and the attendant loss of bowel motility. Constipation may be a symptom of other diseases such as hypothyroidism and diabetes, but after the age of 50 a change in bowel habit suggests the possibility of cancer of the colon or rectum.

Diarrhea is less common than constipation as a long-term complaint. It may result from intolerance to certain foods, notably milk. If milk sugar (lactose) is incompletely digested, it draws water osmotically into the colon and is acted upon there by bacteria to produce gas and lactic acid, which contribute to diarrhea. Among adults who are Asians, blacks, or Ashkenazi Jews, lactose intolerance is found in over 60%. The incidence in populations of northern European descent is only about 15%. Abstention from milk and milk products is curative.

Other food substances may also contribute to chronic diarrhea. Fructose, a sugar found in fruits and used as a sweetener, is inadequately absorbed in some individuals and can lead to diarrhea. The sugar sorbitol is used as a low-calorie substitute for sucrose or glucose because it is poorly absorbed, but diarrhea may result if it is consumed in large quantities.

Diarrhea can, of course, occur as a symptom of serious disease such as regional enteritis or ulcerative colitis. It may also result from infection, particularly by the parasites *Giardia lamblia* or *Entamoeba histolytica*. Bacterial infections caused by *Salmonella*, *Shigella*, and *Campylobacter* may cause short-term diarrhea. Diabetes mellitus, hyperthyroidism, and other systemic diseases can also be associated with diarrhea. *See* BACILLARY DYSENTERY; DIARRHEA; INFANT DIARRHEA.

Irritable bowel syndrome. Among the most common problems in adults is the irritable bowel syndrome. The term is a general one that incorporates such labels as spastic colitis, functional bowel disease, and psychophysiologic gastrointestinal reaction. Population studies have found the incidence of the condition to be as high as 15% in adults; women are more likely to be affected. Anatomical disease is not present, and good health is retained with no long-term consequences. The symptoms include recurrent abdominal pain, constipation or diarrhea that may alternate, and sensation of gaseousness and bloating. The cause has not been identified, but medication and dietary changes, especially the addition of fiber, may offer relief.

Diverticulosis. In diverticulosis, small pouches appear along the wall of the colon. These represent herniations of the mucosa through the muscular layers. In most cases the intestinal musculature is thickened. The disease is more prevalent with increasing age; barium enema radiology or colonoscopy reveals an incidence of 30% or higher among persons in their seventies.

Diverticulosis is associated with high pressures within the colon and particularly its narrowest segment, the sigmoid colon. The pressure pushes the mucosa through the weakest points in the wall, leading to formation of the diverticuli. The symptoms are usually mild, with little or no discomfort or with episodes of pain that are usually experienced in the lower left quadrant of the abdomen. Perforation of the pouches may cause an inflammatory process known as diverticulitis or peridiverticulitis. Erosion of a small artery supplying blood to the colon wall may lead to bleeding.

A high-fiber diet is the standard treatment for diverticulosis, but some authorities also advocate restricting the ingestion of seeds and other foods that might become impacted within a diverticulum. Surgery may be required in cases of peridiverticulitis or hemorrhage.

Cancer. The second most common internal cancer is that of the colon and rectum; more than 130,000 cases are diagnosed each year in the United States.

Cancers frequently arise in benign polyps of the colon, which are stalklike or leafed structures (villous adenomas). A high intake of red meat and low fiber and low vitamin A consumption may be contributing factors; a genetic predisposition has been suggested. Early detection offers the best hope of recovery.

Annual testing for the presence of blood in the stool after age 50 is widely recommended. Detection is possible by means of a barium enema radiological examination or colonoscopy. Early cancers can be removed with a good prognosis; in later stages the chance for recovery is reduced. Small cancers can be removed by colonoscopy; otherwise surgery is required.

Hemorrhoids. Hemorrhoids are veins near the anus that have become distended or occluded. The cause is not fully understood but may be related to straining at stool. Hemorrhoids may be treated with a high-fiber diet, stool softeners, and applications of heat. Severe cases may require surgery.

Appendix. Appendicitis, that is, inflammation of the vermiform appendix, was once common among young people but is now relatively rare; the reason for the sharp drop in frequency is unclear. Appendicitis seems to result when a hardened piece of fecal matter known as a fecalith becomes trapped within the appendix and causes infection that can lead to gangrene. If the condition persists, the appendix may perforate, resulting in widespread infection in the abdomen. Symptoms usually begin with vague discomfort in the midabdomen, occasionally associated with nausea and vomiting; pain and tenderness tend to localize in the right lower quadrant, and then a low-grade fever is likely to develop as well. An inflamed appendix should be removed before it ruptures. *See* APPENDICITIS; CANCER (MEDICINE); DIGESTIVE SYSTEM; TUMOR.

Leonard A. Katz

Bibliography. D. A. Coffman et al., *Gastrointestinal Disorders*, 1987; M. Feldman et al., *Sleisenger and Fordtran's Gastrointestinal Disease*: *Pathophysiology*, *Diagnosis*, *and Management*, 2 vols., 6th ed., 1998; J. Norton, *Gastrointestinal Disorders: A Pathophysiologic Approach*, 1986.

Gastrolith

Any of the pebbles swallowed by animals and retained for a time in the gizzard or stomach, where they serve to grind up the food and in so doing become rounded and highly polished. Birds generally use such pebbles, as do some living reptiles, notably the crocodile and certain lizards. Some of the Mesozoic reptiles also used gastroliths. Articulated skeletons of the Plesiosaurs are generally associated with such highly polished cobbles, even when embedded in formations such as the Niobara chalk, which is otherwise free of gravel. In some instances as many as a half bushel of such stones have been found within the rib cage. Among the dinosaurs the great sauropods likewise used gizzard stones.

Similar highly polished stones found in the arid Southwest of the United States are so abundant and widespread, however, as to indicate that they are due to natural weathering. No satisfactory criteria are known for distinguishing them from gastroliths.

Carl O. Dunbar

Gastropoda

The largest and most varied class in the phylum Mollusca, possibly numbering over 74,000 species and commonly known as snails. *See* MOLLUSCA.

Functional Morphology

The shell is in one piece that, in the majority of forms, grows along a turbinate (equiangular) spiral (**Fig. 1**) but is modified into an open cone in various limpets or is secondarily lost in various slugs.

Effects of torsion. All gastropods, at some time in their phylogeny and at some stage in their development, have undergone torsion. The process does not occur in any other mollusks. It implies that the visceral mass and the mantle shell covering it have become twisted through 180° in relation to the head and foot. As a result of torsion, all internal organs are twisted into a loop. Similarly in gastropods, the mantle cavity (the semi-internal space enclosed by the pallium or mantle) containing the characteristic molluscan gills (ctenidia) has become anterior and placed immediately above and behind the head. The most primitive gastropods, like stem stocks in the other molluscan groups, retain a pair of aspidobranch (bipectinate or featherlike) gills, each with alternating ctenidial leaflets on either side of a ctenidial axis in which run afferent and efferent blood vessels (**Fig. 2**). Lateral cilia on the faces of the leaflets create a respiratory water current (toward the midline and anteriorly) in the direction opposite to the flow of blood through the gills, to create the physiological efficiency of a countercurrent exchange system. *See* COUNTERCURRENT EXCHANGE (BIOLOGY).

As also occurs in other mollusks, the ctenidia form a gill curtain which functionally divides the mantle cavity into an inhalant part (lateral and below) with characteristic osphradia as water-testing sense organs, and an exhalant part into which the anus and renogenital ducts discharge. In zygobranch (two-gilled) gastropods, various shell slits or complex openings have been developed (Fig. 2*c* and *d*) to deal with the problems of sanitation created by the mantle cavity being brought anteriorly above the head. In various ways, these accommodate the exhalant stream of deoxygenated water and the feces and genital and kidney products which accompany it, so that this exhalant discharge is not directly over the head. In the majority of gastropods this problem is resolved by the reduction of the gill pair to a single ctenidium (**Fig. 3**) or to a pectinibranch ctenidium (a one-sided, comb-shaped “half-gill”).

Shell coiling and asymmetry. In a typical snail the shell is turbinate or helicoid, growing as a steadily expanding conic tube along the “screw spiral” of a

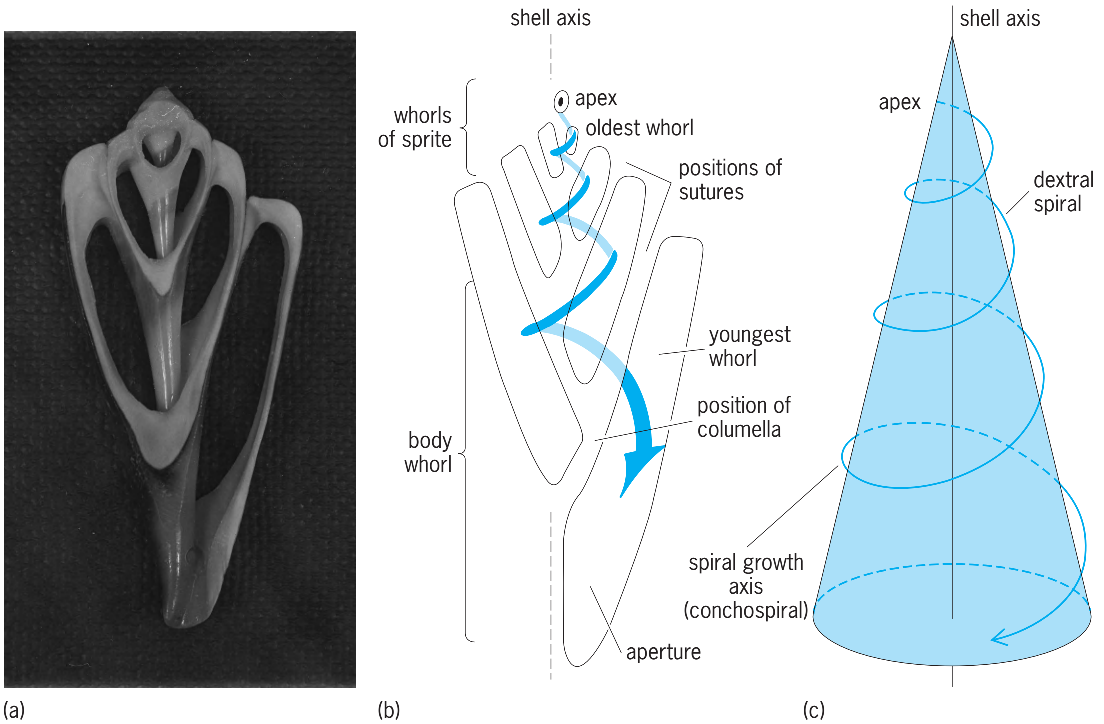


Fig. 1. Typical turbinate shell of gastropods. (*a*) Longitudinal section ground through the shell of a specimen of *Conus purius* to reveal the central columella and spiral of whorls expanding to the aperture. (*b*) Diagram of the shell to show the relationship of the whorls sectioned. (*c*) Stylized representation of the conchospiral line, as viewed with the shell apex tilted toward the observer. This logarithmic or equiangular line is the center line of a steadily expanding conic tube secreted by the mantle edge (at the shell aperture) during growth. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

logarithmic or equiangular line on the surface of a hypothetical cone (Fig. 1). Only a minority of gastropods have planospiral shells or the open limpet form. The majority with turbinate shells cannot have similar left and right halves, and almost all snails are anatomically asymmetric. Even in those gastropods in which the shell is lost and there has been a secondary return to a bilateral symmetry of external features, there remain marked asymmetries of internal anatomy. The majority of turbinate-shelled gastropods have dextral coiling, with a consequent reduction of the pallial organs of the right side, so that the left ctenidium persists. Internal organs are similarly reduced, so that "higher" gastropods have only the left auricle of the heart (the monotocardiac condition), and the left kidney and a single gonad opening to the exterior by the originally right-hand renogenital coelomoduct.

Mantle cavity evolution. From the most primitive forms which retain most clearly the basic effects of torsion (Fig. 2), the further evolution of the gastropods has involved increasing asymmetry of

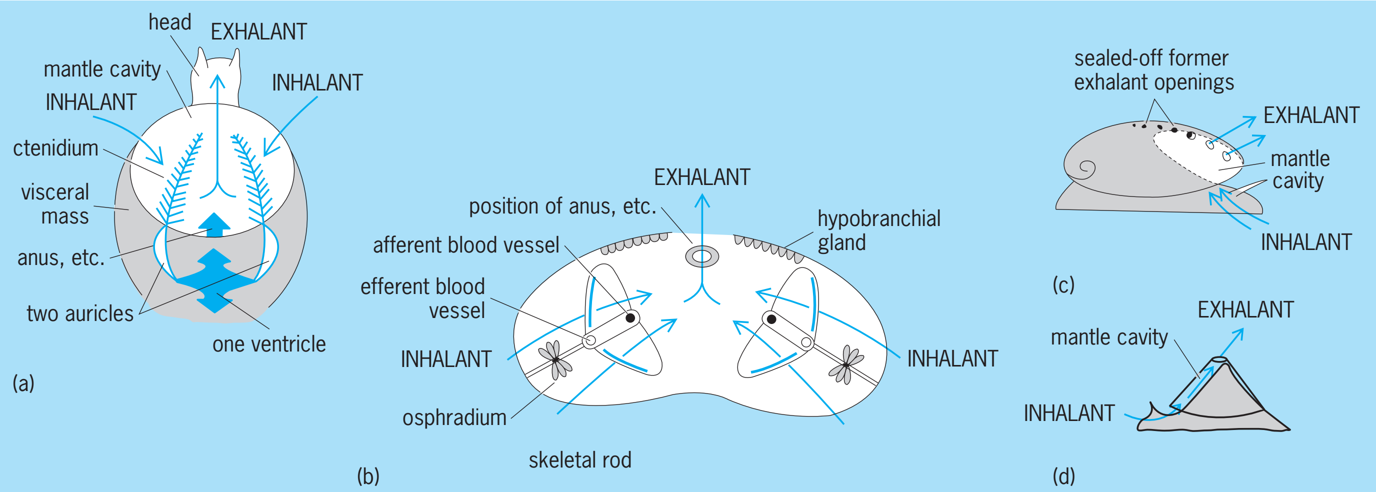


Fig. 2. Mantle cavity and associated organs in a primitive zygobranch gastropod, with a symmetrical pattern of two aspidobranch gills (*a*) in dorsal view and (*b*) in cross section. This arrangement of mantle cavity is found in (*c*) abalones and (*d*) keyhole limpets. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

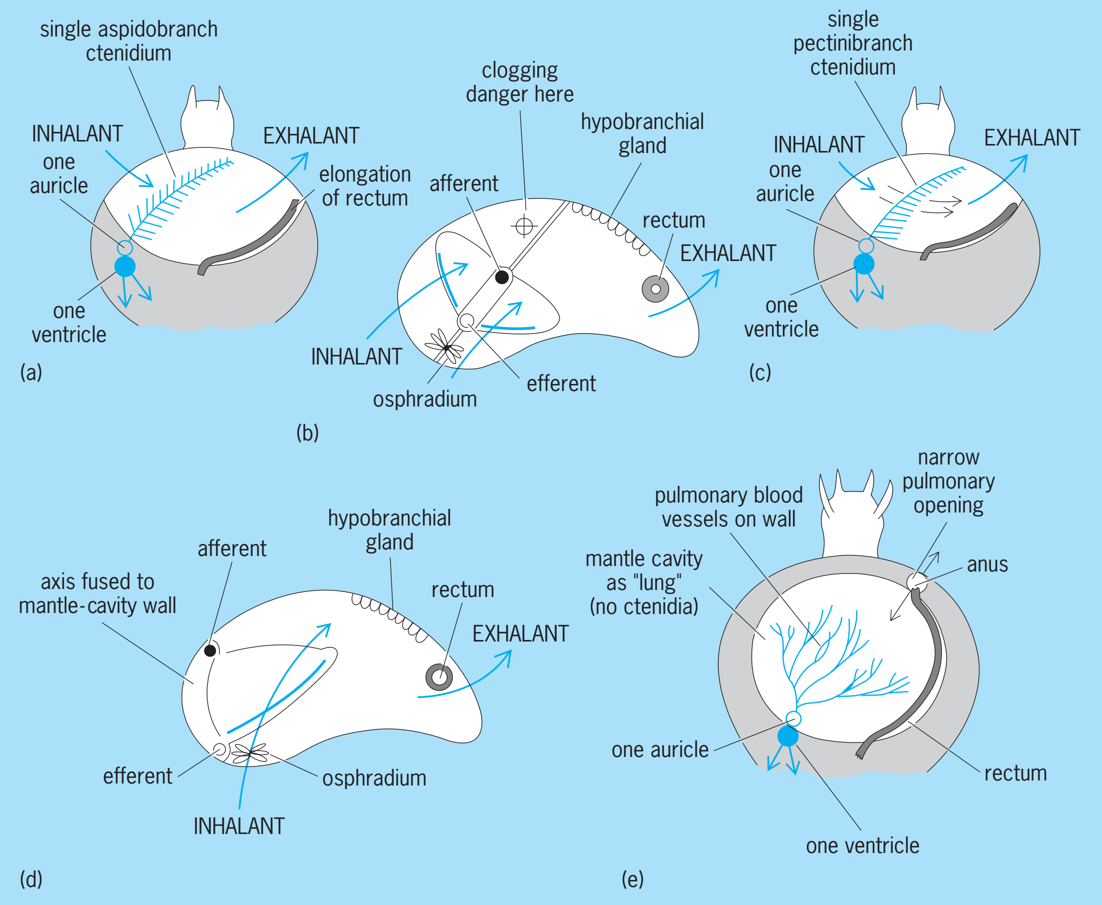


Fig. 3. Further evolution of the mantle cavity in gastropods. The mantle cavity and pallial complex is shown (*a*) in dorsal view and (*b*) in cross section for an asymmetric snail with a single aspidobranch gill, and (*c*, *d*) for an asymmetric snail with a "half-gill" or pectinibranch ctenidium. (*e*) The mantle cavity of a pulmonate land snail in dorsal view shows that there is now no gill and the mantle has become vascularized as a lung wall. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

the mantle cavity organs. Relatively few living gastropods are zygobranch with two ctenidia (and the diotocardiac condition, Fig. 2*a*). These include keyhole limpets such as *Fissurella* and abalones (*Haliotis*).

In most gastropods the right ctenidium is completely lost, and the pallial water flow has become inhalant from the left side of the snail, with the anus and renogenital openings moved over to the right (now exhalant) side (Fig. 3*a* and *b*). A minority of these single-gilled snails retain an aspidobranch gill, with some danger of particulate material clogging the dorsal part of the mantle cavity. Such snails include the limpet genus, *Tectura* (=*Acmaea*), common on both Atlantic and Pacific coasts, the big worldwide group of top shells (*Trochus* and its allies), and the tropical littoral genus *Nerita*. All these forms are ecologically limited to relatively clean water over hard substrata and are unable to invade areas of the sea bottom or seashore covered with mud or silt.

By far the most successful marine gastropods (without such ecological limitation) are those in which the pallial structures are further reduced (Fig. 3*c* and *d*), with a pectinibranch ctenidium whose axis is fused to the mantle wall, resulting in greater hydrodynamic efficiency. Note that this "half-gill" pattern still presents the same functional relationships of ciliary water currents and blood vessels in a countercurrent system. Most familiar snails of the seashore, including such genera as *Busycon* (whelks), *Nassarius* (mud snails), *Littorina* (periwinkles), *Euspira* (=*Polinices*) [moon snails], and very many others, have pectinibranch gills and this highly asymmetric arrangement of the mantle cavity.

A final reduction of pallial structures is found in the pulmonate snails and slugs (Fig. 3*e*), in which the mantle cavity is an air-breathing lung and there are no ctenidia. The other more specialized subclass, the Opisthobranchia, shows detorsion after loss of the shell and a variety of secondary (neomorphic) gills, especially in the extremely beautiful sea slugs or nudibranchs. *See* NUDIBRANCHIA; PULMONATA.

Diversity and Classification

More than half of all molluscan species are gastropods, and they encompass a range from the marine limpets, which can be numbered among the most primitive of all living mollusks, to the highly evolved terrestrial air-breathing slugs and snails. Pulmonates and certain mesogastropod families are the only successful molluscan colonizers of land and freshwaters. The anterior mantle cavity (resulting from torsion), as opposed to the posterior mantle cavity of all other major molluscan stocks, remains diagnostic of the class despite a diversity of functional morphology unequaled by any comparable group in the entire animal kingdom.

Presently classification of gastropods is rather unsettled, even though the phylogeny of this group is being examined vigorously. Consequently, an older arrangement of gastropod taxa, in practice prior to the onset of phylogenetic systematics, is presented here. While most of these taxa are now considered to be artificial, they are useful to consider in that each one demonstrates parallel morphologies and habitats.

The older systematic arrangement of the class Gastropoda involves three somewhat unequal subclasses: Prosobranchia, Opisthobranchia, and Pulmonata.

Prosobranchia. The largest and most diverse subclass is Prosobranchia, which is made up largely of marine snails, all retaining internal evidence of torsion. The prosobranchs have traditionally been divided into three orders: Archaeogastropoda, Mesogastropoda, and Neogastropoda. The primitive Archaeogastropoda are characterized by the presence of one or two aspidobranch (bipectinate) ctenidia, auricles, metanephridia (kidneys), and osphradia (Figs. 2, 3*a* and *b*). Members of the polyphyletic Mesogastropoda, which comprises almost 100 families, have only one pectinobranch (monopectinate) ctenidium (left), auricle, and nephridium (Fig. 3*c* and *d*). Members of the Neogastropoda, which appears to be monophyletic, also exhibit these characters, but their shell has a canal or notch that holds a tubular extension of the mantle (siphon).

In some current classifications, most families of the Mesogastropoda and Neogastropoda (and some Archaeogastropoda) constitute a taxon known as the Caenogastropoda, whereas most of the archaeogastropods are placed within the Patellogastropoda (true limpets) and Vetigastropoda (abalones, top and turban snails, keyhole limpets). *See* MESOGASTROPODA; NEOGASTROPODA; PROSOBRANCHIA.

Opisthobranchia and Pulmonata. The other two subclasses (Opisthobranchia and Pulmonata) are each considerably more uniform than the subclass Prosobranchia and, in both, the effects of torsion are reduced or obscured by secondary processes of development and growth. Even so, in newer classifications they are often combined into a single group (Heterobranchia) along with a few groups (sundials, pyramidellids) formerly classified as prosobranchs.

The marine subclass Opisthobranchia consists largely of sea slugs, in which the shell and mantle cavity are reduced or lost and there is a bilaterally symmetrical adult with a variety of secondary gill conditions. *See* OPISTHOBRANCHIA.

The final subclass, Pulmonata, consists of gastropods with the mantle cavity modified into an air-breathing lung and with no ctenidia (Fig. 3*e*). There are a few littoral marine forms, but the order Basommatophora is mainly made up of the freshwater lung-snails, and the order Stylommatophora consists of the successful land snails such as *Helix* plus a few shell-less families of land slugs. *See* PULMONATA; BASOMMATOPHORA; STYLOMMATOPHORA.

W. D. Russell-Hunter; Robert E. Knowlton

Bibliography. V. Fretter and A. Graham, *British Prosobranch Molluscs*, 2d ed., 1994; F. W. Harrison and A. J. Kohn (eds.), *Microscopic Anatomy of Invertebrates*, vol. 5, 1994, vol. 6B, 1997; R. N. Hughes, *A Functional Biology of Marine Gastropods*, 1986; R. M. Linsley, Shell form and the evolution of gastropods, *Amer. Sci.*, 66:432–441, 1978; W. D. Russell-Hunter, *A Life of Invertebrates*, 1979; J. Taylor (ed.), *Origin and Evolutionary Radiation of the Mollusca*, 1996; E. R. Trueman and M. R. Clarke, *The Mollusca*, vol. 10, 1985.

Gastrotricha

A phylum of minute metazoan animals (sometimes placed in the aschelminth group, in the Cycloneuralia, or in other combinations of phyla), mostly less than a millimeter in length. It presently numbers 600 described species worldwide. Some 380 species have been reported from the marine habitat, with new ones being described every year.

Gastrotrichs comprise two orders, the Macrodasyida and the Chaetonotida. The term Gastrotricha refers to the ventral locomotor cilia by which the animals glide gracefully over the substratum or through its interstices; unlike many other ciliated animals, they cannot move in reverse. Gastrotrichs have a complete digestive tract, with a triradiate muscular sucking pharynx, a simple intestine with a wall only a single cell thick, and a ventral anus. They appear to be selective feeders on bacteria, very small protozoa, and yeasts. The musculature, of obliquely striated fibers, form diverse patterns (including spirals or helices) that reach throughout the body, and allow a wide range of nonlocomotor contortive motion. Most have biflagellate protonephridia, accounting in part for the broad salinity tolerances of many species. The sensory organs are quite sophisticated for an animal of this size. Dual gland adhesive tubes (one with glue, the other with releaser) can occur in groups on the front, side, and rear of the animals; in most chaetonotids these are reduced to a single pair in the rear or entirely lost. The nonchitinous cuticle may have one to many ultrastructural layers; its surface may be naked and translucent or elaborated into thickenings, scales, spines, or hooks (**Figs. 1** and **2**).

Reproduction and development. Gastrotrichs are bisexual, either simultaneously or sequentially male and female. Some macrodasyids pass a spermatophore from donor to recipient animal, with sperm penetrating the body cuticle and moving internally to the eggs; others cross-copulate, some after first moving sperm from a midbody duct to a rear transfer organ. Reproduction in chaetonotids is unique, egg-to-adult ratio being among the greatest in the animal kingdom; a period of asexual reproduction by parthenogenesis is followed by

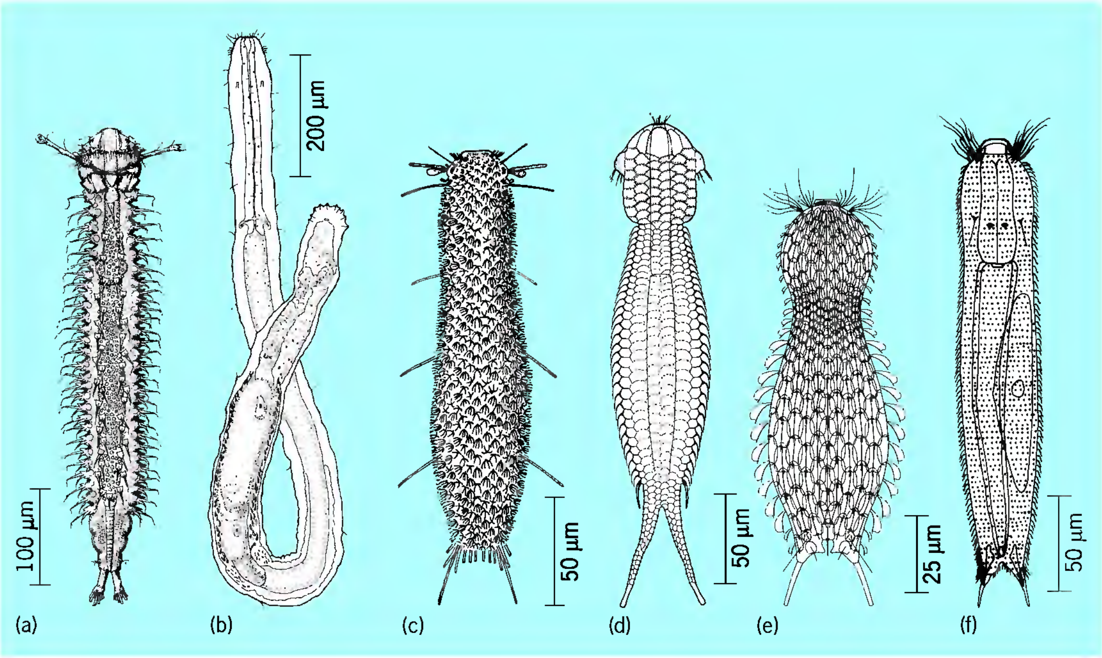


Fig. 1. Diagrams of (a–c) macrodasyid and (d–f) chaetonotid gastrotrichs. (a) *Xenodasys*. (b) *Dolichodasys*. (c) *Thaumastoderma*. (d) *Draculiciteria*. (e) *Halichaetonotus*. (f) *Aspidiophorus*. (Part a after C. Schoepfer-Sterrer, Chordodasys riedli gen. nov., spec. nov., a macrodasyoid gastrotrich with a chordoid organ, Cah. Biol. Mar., 10:391–404, 1969; b after G. D. Gagne, Dolichodasys elongatus n.g., n. sp., a new macrodasyid gastrotrich from New England, Trans. Amer. Microsc. Soc., 96:19–27, 1977; c after A. Remane, Klassen und Ordnungen des Teirreichs, vol. 4(2), Akad. Verlagsges., Leipzig, 1936; d after P. Luporini et al., Contribution à la connaissance des Gastrotriches des côtes de Toscane, Cah. Biol. Mar., 12:433–455, 1971; e after H. Mock, Chaetonotoidea (Gastrotricha) der Nordseeinsel Sylt, Mikrofauna Meersebodens, 78:1–107, 1979; f after W. D. Hummon, Gastrotricha from Beaufort, North Carolina, U.S.A., Cah. Biol. Mar., 15:431–446, 1974).

cross-fertilization between reciprocal hermaphrodites. Eggs then rupture through the body cuticle or exit through a ventral pore and are deposited on the surface or in crevices of substratum particles. Reproductive biology and RNA phylogenies are areas of active research.

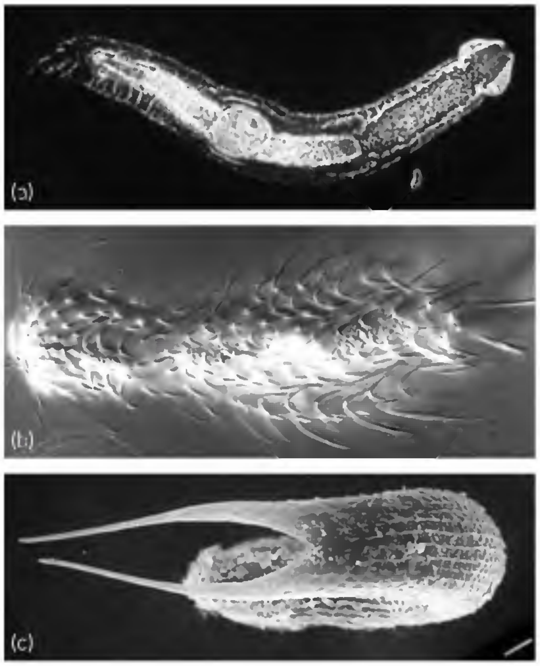


Fig. 2. Photographic images of (a) *Cephalodasys littoralis*, 670 m, via phase contrast microscopy; (b) *Chaetonotus siciliensis*, 300 m, via differential interference contrast microscopy; and (c) *Musellifer sublitoralis*, 330 m, via scanning electron microscopy, (scalebar 10 m). (Parts a and c after W. D. Hummon, c after M. Balsamo)

Most gastrotrich eggs are elliptical and measure 40 × 60 micrometers; those of macrodasyids are smooth, of chaetonotids mostly rough or spiny. Usually, only one egg is laid in each clutch. Cleavage is modified radial with blastomeres being subequal in size. Embryogenesis at 20°C (68°F) takes 10–12 days for macrodasyids or 1–3 days for chaetonotids. Development is direct: juveniles being about 100 μm in length and growing to adulthood in 30 days, in macrodasyids, or 2–5 days, in chaetonotids. The life span of macrodasyids can be up to 6 months, probably averaging less than a month, for chaetonotids up to a month, averaging less than a week; populations can increase rapidly under suitable conditions. Chaetonotids have fewer than a thousand cells, among the least among metazoans; they are cell-constant and show eutely (all adults have the same number of cells). Chaetonotids may lay eggs that resist temperature extremes and drying, some being able to hatch even after a decade. *See* CELL CONSTANCY.

Biogeography and evolution. Gastrotrichs appear to be regionally cosmopolitan, with 20–30% having broad distributions within continents, and 10–15% between continents; endemism probably does not exceed 20%. Biogeographical research has broadened from Europe, India, and North America to include South America, Australia, and the Far East.

In having species with monociliated epithelial cells, the phylum Gastrotricha is among the most primitive in the aschelminth group of phyla.

Gastrotrichs and nematodes probably share a common ancestor, which in turn was descended from a stock that included gnathostomulids and turbellarianoid animals. *See* CHAETONOTIDA; GNATHOSTOMULIDA; MACRODASYIDA; NEMATA; ROTIFERA; TURBELLARIA.

William D. Hummon

Bibliography. K. G. Adiyodi and R. G. Adiyodi (eds.), *Reproductive Biology of Invertebrates*, 6 vols., 1983–1993; R. P. Higgins and H. Thiel (eds.), *Introduction to the Study of Meiofauna*, 1988; C. Nielsen, *Animal Evolution: Interrelationships of the Living Phyla*, 2d ed., 2001; S. D. Rundle, A. L. Robertson, and J. M. Schmid-Araya (eds.), *Freshwater Meiofauna: Biology and Ecology*, 2002; E. E. Ruppert, Gastrotricha, in F. W. Harrison (ed.), *Microscopic Anatomy of Invertebrates*, vol. 4, 1991; W. H. Thorp and A. P. Covich (eds.), *Ecology and Systematics of North American Freshwater Invertebrates*, 1991.

Gastrulation

A set of cellular and epithelial movements that transform the blastula stage of embryonic development into a more complex arrangement of three germ layers. The outer germ layer, which will be exposed to the external environment, is the ectoderm. It gives rise to the epidermis and nervous system of the developing organism. The inner layer of cells, the endoderm, forms the inner lining of the digestive tract and its derivatives. An intermediate layer, the mesoderm, develops into a diverse array of structures, including the skeleton, muscles, kidneys, and circulatory system. The ectoderm and endoderm are usually epithelial, consisting of closely packed sheets of cells connected by tight junctions. The mesoderm is sometimes epithelial, but at other times it is mesenchymal, forming a loose arrangement of cells surrounded by extracellular material. The three germ layers are formed in nearly all animal embryos, while the movements that generate them vary. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY); GERM LAYERS.

Beginning of morphogenesis. Gastrulation marks the onset of dramatic cellular and epithelial movements that are collectively called morphogenesis because they generate the embryo's body plan. During morphogenesis, individual cells move relative to each other, while entire epithelia spread, fold, and change shape. The amazing complexity and order of these movements are best appreciated in time-lapse movies, which show how cells divide, wedge themselves between other cells, and change between epithelial and mesenchymal states. These cellular behaviors are based on the orchestrated synthesis of proteins that control the shape, adhesion, and movements of cells. *See* EMBRYONIC INDUCTION; MORPHOGENESIS.

Gastrulation in sea urchins. Sea urchins illustrate gastrulation with almost schematic clarity. The sea urchin blastula is a sphere of epithelial cells surrounding a blastocoel filled with a gelatinous fluid (**Fig. 1***a*). A group of cells with long cilia marks the blastula's animal pole; the opposite pole is the vegetal pole. Gastrulation begins when the vegetal cells become tall and columnar, forming the vegetal plate (Fig. 1*b*). Next, some cells of the vegetal plate begin a morphogenetic movement known as ingression: They break the epithelial junctions that have sealed them to neighboring cells and move into the blastocoel as single cells (Fig. 1*c*). Called primary mesenchyme cells, they later form the skeleton of spicules supporting the arms of the sea urchin larva.

A major event in gastrulation is the formation of the primitive gut, or archenteron. It begins

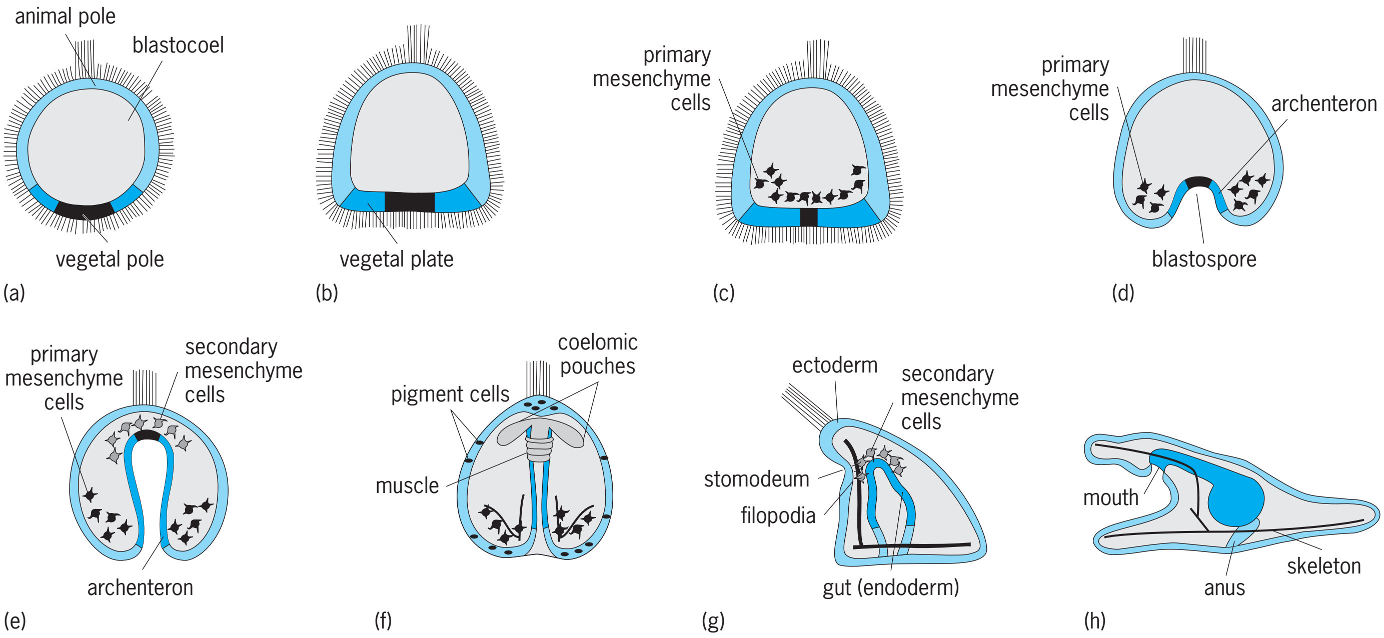


Fig. 1. Sea urchin gastrulation. (*Reprinted with permission from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

with a common type of epithelial movement called invagination, in which an epithelium buckles inward to create a depression (Fig. 1*d*). The opening of the archenteron, called the blastopore, marks the site where endodermal and mesodermal cells move inside the embryo. The short archenteron rudiment then undergoes another typical epithelial movement known as convergent extension, that is, an elongation in one dimension with simultaneous shrinkage in one or two other dimensions (Fig. 1*e*). This is achieved by two cellular activities. First, individual archenteron cells elongate in a coordinate way. Second, the same cells slide past one another, so that the number of cells counted along the archenteron's length increases while the number of cells in its circumference decreases.

While archenteron elongation is still underway, a set of secondary mesenchyme cells become motile at the archenteron tip (Fig. 1*e*). Some of them form pouches of mesodermal epithelium, which displace the blastocoel by generating a secondary body cavity called the coelom (Fig. 1*f*). Other secondary mesenchyme cells give rise to pigment cells, muscles, and blood cell equivalents. While still attached to the archenteron tip, secondary mesenchyme cells also initiate the final phase of archenteron elongation by extending thin cytoplasmic processes called filopodia. They probe the inside of the blastocoel wall, attach at suitable sites and then contract, orienting the archenteron tip toward the ventral side of the embryo (Fig. 1*g*). Here, the archenteron tip is met by the stomodeum, a depression in the oral surface of the ectoderm. Fusion of the stomodeum with the archenteron tip generates the mouth of the larva, while the blastopore becomes the anus (Fig. 1*h*).

Gastrulation in humans. In humans, the inner cell mass of the blastocyst forms a disc consisting of two layers called epiblast and hypoblast. In a morphogenetic movement called delamination, the epiblast splits into the embryonic epiblast and the amnion (**Fig. 2**). The amniotic cavity between them will expand and contain the amniotic fluid that surrounds the growing fetus until birth. The following description focuses on the embryonic epiblast, called simply epiblast from here on, which will give rise to the embryo proper.

Gastrulation begins with formation of the primitive streak, a pair of ridges with a groove between them, extending along the epiblast's midline (**Fig. 3**). The thickened anterior end of the streak is known as Hensen's node (like the corresponding structure in the chicken embryo, from which much of our knowledge on human gastrulation has been gleaned). The primitive streak becomes the major gastrulation site, similar to the blastopore in the sea urchin embryo, where future endoderm and mesoderm cells move inside. In humans, epiblast cells converge to the primitive streak. Here they undergo a morphogenetic movement known as involution, which is defined as an inward movement of an expanding epithelium around an edge. As soon as the epiblast cells have involuted around the ridges of the primitive streak and disappeared into the groove, they become mesenchymal and ingress into the space between the epiblast and hypoblast (**Fig. 4**).

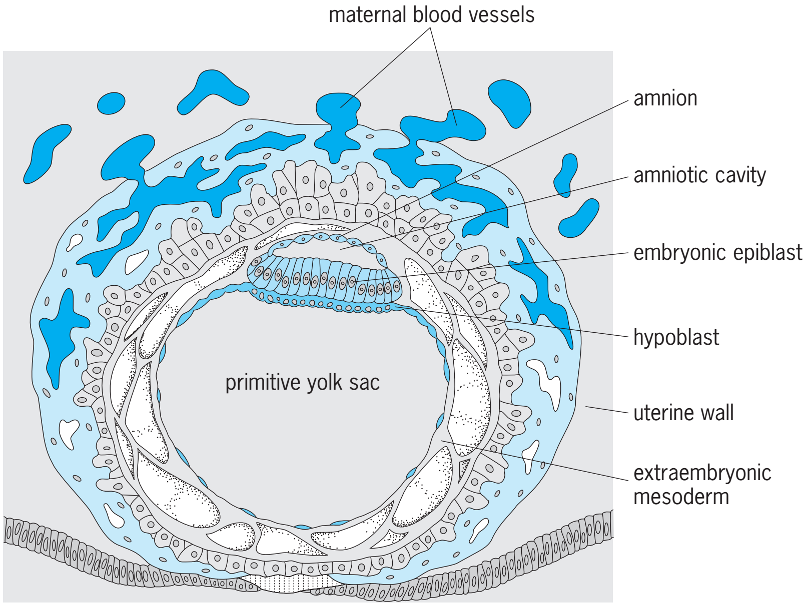


Fig. 2. Human embryo at day 12. (*Reprinted with permission from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

The fate of the ingressing cells depends on the time of ingression and the direction of their further movement. The cells that ingress first form the embryonic endoderm, which will line the inside of the digestive tract. Subsequently, Hensen's node recedes to the posterior, thus shortening the primitive streak. Later, ingressing cells that stay along the midline, in the wake of the receding Hensen's node, become the axial mesoderm, which will form the precursor of the embryonic backbone. Cells ingressing to more lateral positions form paraxial and lateral mesoderm, which make trunk musculature, kidneys, blood vessels, heart, and blood. Epiblast cells that do not involute remain on the outside and form the ectoderm,

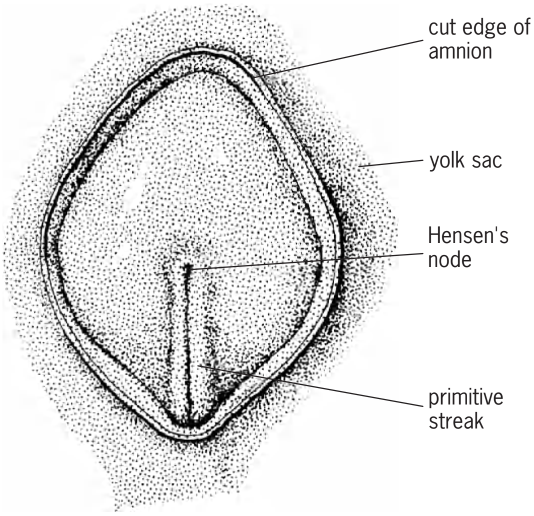


Fig. 3. Human epiblast in dorsal view at day 16. The amnion has been cut away. (*Modified with permission from B. M. Carlson, Human Embryology and Developmental Biology, Mosby, 1994*)

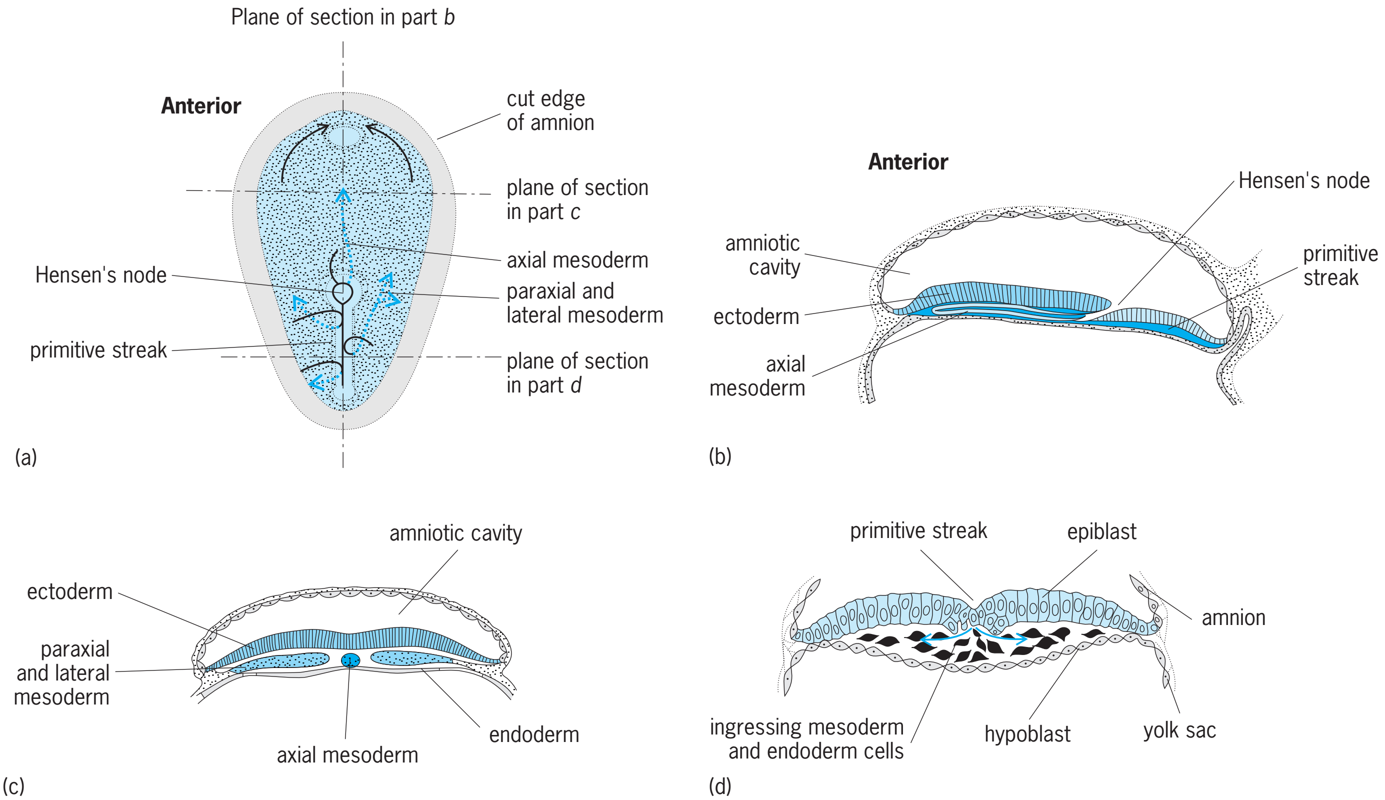


Fig. 4. Gastrulation in the human embryo. (*Reprinted with permission from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

which gives rise to the epidermis, brain, and spinal cord.

In broad terms, the sea urchin gastrula is a three-layered sphere, with ectoderm outside, endoderm inside, and mesoderm between. In the human, the same germ layers form a triple disk that will be transformed into a cylindrical embryo by subsequent morphogenetic movements.

Klaus Kalthoff

Bibliography. B. M. Carlson, *Human Embryology and Developmental Biology*, Mosby, 1994; K. Kalthoff, *Analysis of Biological Development*, 2d ed., McGraw-Hill, Boston, 2001.

Gate circuit

An electronic circuit that consists of elements, which may be transistors, diodes, or resistors, combined in such a manner that they perform a logic operation. Gate circuits are the most basic building blocks of a digital system. These circuits have one or more inputs and one output which is a boolean function of the inputs. The input and output signals can have only two discrete values, low (for example, 0 V) and high (for example, 3.3 V). These values are usually represented as 0 and 1, or "false" and "true," respectively.

Whereas the early gate circuits consisted of diodes, resistors, and transistors, the majority of gate circuits nowadays are built exclusively with transistors. The dominant technology for fabricating gate circuits is the metal-oxide-semiconductor (MOS) silicon method, followed by the silicon bipolar and gallium arsenide (GaAs) techniques. The manufacturing process has become so sophisticated that transistors smaller than 1 square micrometer can be fabricated, allowing the placement of millions of gate circuits on a silicon chip the size of a fingernail. The main advantage of MOS technology is that it gives rise to very low power circuits that can still operate at relatively high clock speed. It is these characteristics

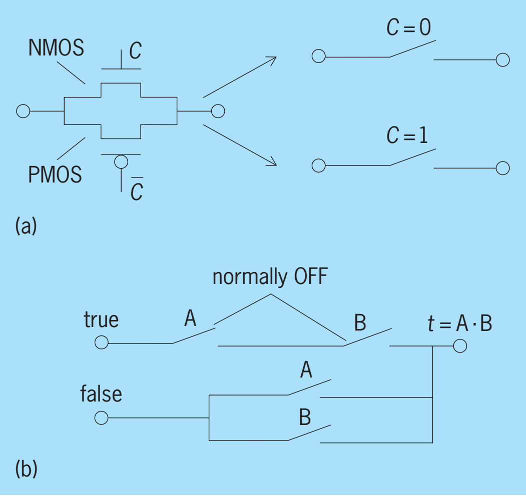


Fig. 1. Switch circuit architecture. (*a*) CMOS transmission gate, with an NMOS and a PMOS transistor in parallel, clocked at opposite clock signals (*C* and $\bar{C}$). (*b*) Switch circuit for two-input logic AND function, with two normally-OFF switches in series and two normally-ON switches in parallel. Each switch represents a transmission gate.

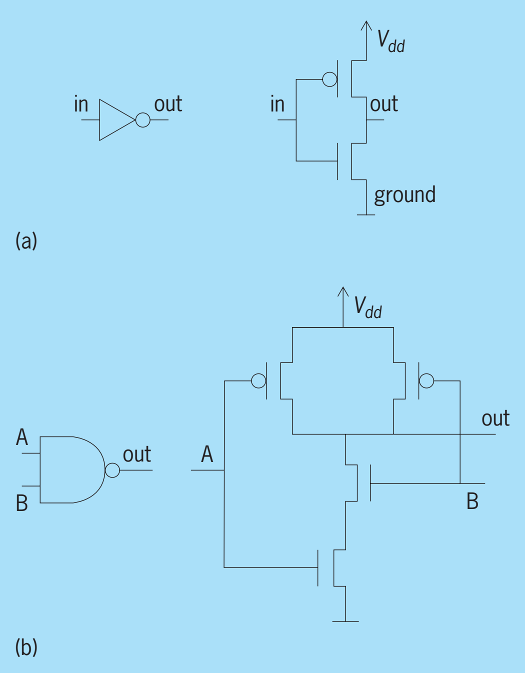


Fig. 2. Logic symbols and corresponding gate circuits. (*a*) CMOS inverter circuit built with a PMOS and an NMOS transistor. (*b*) Two-input NAND circuit.

that have allowed the fabrication of very complex digital systems such as microprocessors and memories. *See* INTEGRATED CIRCUITS; MICROPROCESSOR; SEMICONDUCTOR MEMORIES.

The transistors in a gate circuit are used as ON-OFF switches. By combining these transistors in a certain way, it is possible to realize logical, arithmetical, and memory functions. There are two type of MOS transistors: NMOS and PMOS field-effect transistors (FETs), corresponding to a normally-OFF or normally-ON switch, respectively. Circuits in which both types of transistors are used are called CMOS (complementary MOS) circuits. CMOS circuits now constitute the majority of gate and logic circuits.

Transmission gate or switch circuits. In switch circuits an MOS transistor is used to pass or block the flow of information in a similar fashion to a mechanical switch. By placing these switches in a network, it is possible to realize different logic functions. A transistor used in this fashion is often called a pass-transistor. In order to improve the switching characteristics, an NMOS switch and a PMOS switch are placed in parallel, each clocked at opposite clock signals. Such a combination is called a CMOS transmission gate (**Fig. 1***a*). Several transmission gates can be combined to form a logic AND circuit (Fig. 1*b*). Each switch represents a transmission gate composed of two transistors in parallel (Fig. 1*a*). The two switches in series are normally-OFF, while the two switches in parallel are normally-ON switches. The network of switches (Fig. 1*b*) is such that when both inputs A and B are true the two series switches are ON and will pass the information. As soon as one of the inputs is false, the corresponding series switch will be OFF and the parallel switch will be ON, making the output false, thus implementing an AND function. Transmission gates are also widely used to implement more complex circuits such as delay elements, flip-flops, and other transmission gate logic structures. *See* SWITCHING CIRCUIT.

Logic gates. An alternative way to realize logic functions is to make use of logic gates. The simplest gate circuit, the inverter, takes an input signal and presents the inverted signal at the output (**Fig. 2***a*). This circuit works as follows: When 0 V (corresponding to a logic 0 or "false" signal) is applied to the input, the NMOS transistor (bottom one) will be OFF, while the PMOS transistor will be ON, connecting the output node to the voltage V_{dd} (corresponding to a logic 1). In contrast, a logic 1 (that is, V_{dd}) applied to the input will switch the NMOS transistor ON and the PMOS OFF, thus connecting the output node to the ground (logic 0).

One of the two transistors in this circuit will always be OFF when the output is stable. This implies that there is no current flowing in the transistors so that the quiescent power dissipation goes to zero. Only when the transistors are switching from the ON to the OFF state will both be ON for a short time, giving rise to dynamic power dissipation. This is one of the main advantages of CMOS circuits and is the reason why the majority of integrated circuits are fabricated in CMOS.

An example of a NAND circuit is shown in Fig. 2*b*. The operation of this circuit can be easily understood. Since the NMOS transistors are normally-OFF and the PMOS transistors are normally-ON, the output will be connected to V_{dd} (logic 1) as soon as one of the input signals is low because at least one of the PMOS transistors will be ON. Only when both inputs are high will the two NMOS transistors be ON, thus making the output low. By using additional transistors, it is possible to implement more complex functions such as arithmetic or memory elements.

The above examples are all CMOS gate circuits. Similar logic functions can be realized with bipolar or

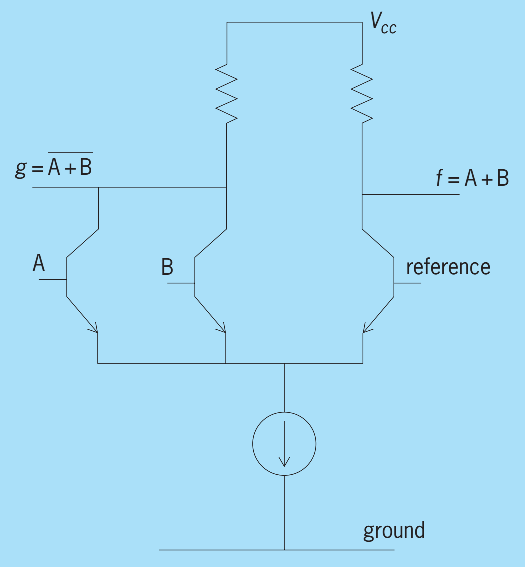


Fig. 3. Emitter-coupled logic (ECL) two-input OR/NOR circuit, consisting of a differential pair.

gallium arsenide transistors. The transistor-transistor logic (TTL) circuits were very popular before the CMOS technology came to maturity. An alternative bipolar technology is the emitter-coupled logic (ECL) circuits. These are used mainly to implement very high speed logic circuits such as those found in the core of processors of supercomputers. A two-input OR/NOR emitter-coupled logic circuit (**Fig. 3**) consists of a differential pair. As soon as one of the inputs A or B is high, the output g will go low and f will go high, implementing the OR and NOR functions, respectively. *See* LOGIC CIRCUITS; TRANSISTOR.

Jan Van der Spiegel

Bibliography. S.-M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits*, 2d ed., McGraw-Hill, New York, 1998; J. Rabaey, *Digital Integrated Circuits*, Prentice Hall, Upper Saddle River, NJ, 1996.

Gauge theory

The theoretical foundation of the four fundamental forces of nature, the electromagnetic, weak, strong, and gravitational interactions. Gauge symmetry lies at the heart of gauge theory. A gauge symmetry differs from an ordinary symmetry in two important respects:

1. Gauge symmetry is a local symmetry rather than a global symmetry. For a local symmetry, the element of the symmetry group (G) that acts on the fields of a theory at a space-time point ($\mathbf{x}$, t) depends on the position $\mathbf{x}$ and time t, whereas for a global symmetry a fixed group element acts on fields at different space-time points.

2. A gauge transformation leaves a physical state invariant. Gauge symmetry reflects a redundancy in the variables used to describe a physical state. By contrast, a global symmetry acting on a physical state in general produces a new, distinct physical state. *See* FUNDAMENTAL INTERACTIONS; SPACE-TIME.

Electromagnetism. The simplest example of a gauge theory is electromagnetism. In classical electrodynamics, gauge invariance reflects the arbitrariness that exists in choosing the potentials $\mathbf{A}(\mathbf{x}, t)$ and $\phi(\mathbf{x}, t)$ to represent the electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$, according to Eqs. (1), where c is the

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) &= -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{B}(\mathbf{x}, t) &= \nabla \times \mathbf{A} \end{aligned} \tag{1}$$

speed of light. If $\Lambda(\mathbf{x}, t)$ is an arbitrary scalar field, Eqs. (2) define a gauge transformation. The potentials $\mathbf{A}'$ and ϕ' may equally well be used to represent the electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$. By using units such that $c = 1$, a more compact, manifestly Lorentz covariant notation may be adopted, in which there is a single 4-vector potential $A_\mu = (\phi, -\mathbf{A})$, so that the field strengths are components of a single antisymmetric, second-rank 4-tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, whose components are given by Eq. (3).

$$\begin{aligned} \phi \to \phi' &= \phi - \frac{1}{c}\frac{\partial \Lambda}{\partial t} \\ \mathbf{A} \to \mathbf{A}' &= \mathbf{A} + \nabla\Lambda \end{aligned} \tag{2}$$

$$\begin{aligned} F_{\mu\nu} &= \begin{pmatrix} F_{00} & F_{01} & F_{02} & F_{03} \\ F_{10} & F_{11} & F_{12} & F_{13} \\ F_{20} & F_{21} & F_{22} & F_{23} \\ F_{30} & F_{31} & F_{32} & F_{33} \end{pmatrix} \\ &= \begin{pmatrix} 0 & +E_x & +E_y & +E_z \\ -E_x & 0 & -B_z & +B_y \\ -E_y & +B_z & 0 & -B_x \\ -E_z & -B_y & +B_x & 0 \end{pmatrix} \end{aligned} \tag{3}$$

See MAXWELL'S EQUATIONS; POTENTIALS; RELATIVISTIC ELECTRODYNAMICS.

In nonrelativistic quantum mechanics, gauge invariance is realized as follows. The Schrödinger equation for a particle with an electromagnetic charge q and mass m is Eq. (4), where $\hat{H}$ is the hamiltonian operator, $\hat{p}$ is the momentum operator, Ψ is the wave function of the particle, and $\hbar$ is Planck's constant divided by 2π. Equation (4) is invariant under the gauge transformation given by Eqs. (5).

$$\begin{aligned} \hat{H}\Psi &= \frac{1}{2m}\left(\hat{\mathbf{p}} - \frac{q}{c}\mathbf{A}\right)^2 \Psi + q\phi\Psi \\ &= \frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}\mathbf{A}\right)^2 \Psi + q\phi\Psi \\ &= i\hbar\frac{\partial}{\partial t}\Psi \end{aligned} \tag{4}$$

$$\begin{aligned} \Psi \to \Psi' &= \exp\left[\frac{iq\Lambda}{\hbar c}\right] \cdot \Psi \\ \phi \to \phi' &= \phi - \frac{1}{c}\frac{\partial \Lambda}{\partial t} \\ \mathbf{A} \to \mathbf{A}' &= \mathbf{A} + \nabla\Lambda \end{aligned} \tag{5}$$

Electromagnetism is a $U(1)$ gauge symmetry [where U(1) is the one-dimensional unitary group, which is represented by the complex numbers $e^{i\varphi}$ with $0 \leq \varphi < 2\pi$] because a gauge transformation rotates the phase of the wave function in a space-time–dependent manner and adjusts the potentials $\mathbf{A}$ and ϕ accordingly. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION; UNITARY SYMMETRY.

In classical electrodynamics, gauge invariance plays a more formal role, because classically it is possible to formulate electromagnetism entirely in terms of the gauge-invariant, physically observable field strengths $\mathbf{E}$ and $\mathbf{B}$ without recourse to the potentials $\mathbf{A}$ and ϕ. Although at the classical level charged particles interact only with the field strengths $\mathbf{E}$ and $\mathbf{B}$ through the Lorentz-force equation, this is not true quantum mechanically. Any quantum-mechanical description of the electromagnetic interaction of charged particles must involve the potentials $\mathbf{A}$ and ϕ, as illustrated by the Aharonov-Bohm effect. *See* AHARONOV-BOHM EFFECT.

Geometric interpretation. Gauge invariance may also be regarded in a more geometric way by using the language of connections and fiber bundles. Although the ordinary derivatives of the wave function

$\nabla\psi$ and $\partial_t\psi$ are not gauge invariant, the covariant derivatives, defined by Eqs. (6), acting on the wave

$$\begin{aligned} \mathbf{D} &= \nabla - \frac{iq}{\hbar c}\mathbf{A} \\ D_t &= \partial_t + \frac{iq}{\hbar}\phi \end{aligned} \tag{6}$$

function $\psi(\mathbf{x}, t)$ give gauge-invariant quantities. The idea of a covariant derivative is based on the notation of parallel transport. In electromagnetism the gauge field $A_\mu(x)$ defines a parallel transport for quantum phases (that is, how quantum phases at different space-time points should be compared). If C is a path connecting the space-time point x to the point y, then the quantity $U(x, y; C)$ defined by Eq. (7) is

$$U(x, y; C) = \exp\left(\frac{iq}{\hbar c}\int_{C_{x\to y}} dx^\mu A_\mu\right) \tag{7}$$

the operator used to parallel-transport the phase of a field of charge q from x to y along the path C. (In this and subsequent equations, whenever an index such as μ is repeated, it is to be summed over its possible values, in this case 0, 1, 2, 3.) The expression given by Eq. (8) is then gauge invariant. In general,

$$\Psi(y)^\dagger U(x \to y;\ C)\Psi(x) \tag{8}$$

the parallel transport depends on the choice of path C connecting the end points. This path dependence leads to the notion of curvature. Upon parallel transport around a closed path C, starting and ending at x_0, the wave function acquires a phase given by Eq. (9),

$$\begin{aligned} U &= \exp\left[\frac{iq}{\hbar c}\oint_C dx^\mu A_\mu\right] \\ &= \exp\left[\frac{iq}{\hbar c}\int_S dS^{\mu\nu}F_{\mu\nu}\right] \end{aligned} \tag{9}$$

where S is a two-dimensional surface in space-time whose boundary ∂S is C. The antisymmetric tensor $F_{\mu\nu}$, also known as the gauge curvature, is the electromagnetic field strength. On an infinitesimal level, the gauge curvature reflects the lack of commutation of the covariant derivatives, according to Eq. (10). Here,

$$F_{\mu\nu} = \frac{-i}{(q/\hbar c)}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{10}$$

$[D_\mu, D_\nu]$ stands for $D_\mu D_\nu - D_\nu D_\mu$, called the commutator of D_μ and D_ν. *See* FIBER BUNDLE; RIEMANNIAN GEOMETRY.

Quantum electrodynamics. The action for an electromagnetic field coupled to a free scalar field is given by Eq. (11). Here $D_\mu = (\partial_\mu + ieA_\mu)$, where

$$S = \int d^4x\left[\frac{-1}{4}F_{\mu\nu}F_{\mu\nu} + [D_\mu\Psi]^\dagger[D^\mu\Psi]\right] \tag{11}$$

e is the charge of the electron and units are chosen such that $\hbar = c = 1$. When the fields $A_\mu(x)$ and $\phi(x)$ are taken to be quantum fields rather than classical fields, the resulting theory is called quantum electrodynamics (QED). This theory cannot be solved exactly; however, QED may be formulated perturbatively, as a power series in the fine-structure constant $\alpha = e^2/\hbar c$. Originally, perturbative QED was fraught with difficulties, because higher powers in α gave infinities. A successful program of regularization and renormalization (to render finite these infinities and to relate the bare parameters appearing in the action to physically observable quantities) was developed independently by R. Feynman, J. Schwinger, and S. Tomonaga in the late 1940s, making QED a theory consistent to all orders of perturbation theory. *See* ACTION; QUANTUM ELECTRODYNAMICS; RENORMALIZATION.

One effect of higher-order quantum corrections in QED is vacuum polarization, which causes running of the coupling constant α. At the classical level (to lowest order in α) the strength of the electromagnetic interaction is the same at all length scales (that is, the inverse square law is valid). However, radiative corrections cause α to increase as increasingly smaller length scales are probed, because of vacuum polarization. Heuristically, this may be understood as screening. On length scales shorter than the electron Compton wavelength, a polarized cloud of virtual electron-positron pairs shields bare charge because of the strong electric field, in much the same way as a charge in a dielectric medium is partially shielded because of the polarization of the dielectric medium. On smaller scales, more charge is seen, so that the effective strength of the coupling seems to increase.

Nonabelian gauge theory. The concept of a gauge theory may be generalized to larger, nonabelian Lie groups, such as $G = \mathrm{SU}(2)$, SU(3), SU(5), or SO(10). A nonabelian Lie group G has generators τ_a $(a = 1, \ldots, d)$, with the commutation relations $[\tau_a, \tau_b] = f^c_{ab}\tau_c$. A function $\Omega(x)$, which depends on the space-time point x and has values in the group G, will be used to generate a gauge transformation. This function can be written in the form $\Omega(x) = \exp[i\omega^a(x)\tau_a]$, at least locally, where the quantities ω^a are generalizations of the quantity Λ introduced in electrodynamics. The theory has matter fields, transforming according to the representations $\Gamma^{(i)}$ of G, respectively. Each representation consists of a set of hermitian matrices satisfying the commutation relations $[\Gamma^{(i)}_a, \Gamma^{(i)}_b] = f^c_{ab}\Gamma^{(i)}_c$. The gauge field A_μ takes values in the Lie algebra and may be expanded in components as $A_\mu = A^a_\mu\tau_a$. Under a gauge transformation, these fields transform according to Eqs. (12), where the

$$\begin{aligned} &\Psi^{(i)} \to \Psi^{(i)\prime} = \Omega^{(i)}\Psi^{(i)} \\ &A_\mu \to A'_\mu = \Omega^{(a)}A_\mu[\Omega]^{(a)-1} + i[\Omega]^{(a)-1}\left(\partial_\mu\Omega^{(a)}\right) \end{aligned} \tag{12}$$

superscript (a) indicates the adjoint representation. The covariant derivative is $D_\mu = (\partial_\mu + igA^a_\mu\Gamma_a)$, where the representation of Γ_a corresponds to that of the field on which D_μ acts. The gauge curvature is given by Eq. (13), and transforms under gauge transformations according to Eq. (14).

$$F^a_{\mu\nu} = -[D_\mu, D_\nu]^a = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a_{bc} A^b_\mu A^c_\nu \quad (13)$$

$$F_{\mu\nu} \rightarrow F'_{\mu\nu} = \Omega^{(a)} F_{\mu\nu} [\Omega^{(a)}]^{-1} \quad (14)$$

It is customary to absorb the coupling constant g into a rescaling of the gauge field, so that the action has the form given by Eq. (15), where tr $[F_{\mu\nu}\, F^{\mu\nu}]$ is

$$S = \int d^4x \left[\frac{-1}{4g^2} \mathrm{tr} \left[F_{\mu\nu} F^{\mu\nu} \right] + \sum_{(i)} \left(D_\mu \Psi^{(i)} \right)^\dagger \left(D_\mu \Psi^{(i)} \right) \right] \quad (15)$$

the trace of the product of the matrices $F_{\mu\nu}$ and $F^{\mu\nu}$. *See* GROUP THEORY; LIE GROUP; MATRIX THEORY.

Electroweak and strong interactions. In 1954 C. N. Yang and R. L. Mills suggested gauging the SU(2) isospin symmetry, thus developing the first nonabelian gauge theory, also known as Yang-Mills theory. In 1971 G. 't Hooft demonstrated the renormalizability of nonabelian gauge theory. Nonabelian gauge theory is the foundation of the electroweak and strong interactions. In the electroweak theory, formulated by S. Weinberg and A. Salam in 1967, the gauge group $G = \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ is spontaneously broken to $\mathrm{U}(1)_Q$, the gauge group of ordinary electromagnetism, by the condensation of a Higgs field. There are four kinds of particles that mediate the gauge interactions, called gauge bosons: two massive charged weak gauge bosons, the W^+ and W^-, with a mass of 80.6 GeV; a neutral gauge boson, the Z, with a mass of 91.2 GeV; and finally a massless photon (γ). The $W^\pm$ and the Z are responsible for the charged and neutral weak interactions, respectively, and the photon is the gauge particle responsible for the electromagnetic interaction. In quantum chromodynamics (QCD) the symmetry group is $\mathrm{SU}(3)_{\mathrm{color}}$, which remains unbroken. There are eight gauge bosons called gluons. Quarks, which come in three colors, carry a color charge, and the quark field forms an SU(3) triplet. *See* ELECTROWEAK INTERACTION; SYMMETRY BREAKING.

Asymptotic freedom and confinement. One remarkable property of quantum chromodynamics is that color charge is antiscreened rather than screened, a feature arising from the nonabelian character of the gauge symmetry. This situation is opposite to that in QED. Consequently, as shorter distance scales are probed, the coupling constant α_{QCD} decreases, so that at very short distances QCD approaches a free (noninteracting) field theory. This feature, pointed out in 1973 by D. Gross and F. Wilczek and by H. D. Politzer, is known as asymptotic freedom and provides the theoretical foundation for the parton model. Although on large scales QCD is a strongly coupled theory, at very small scales, which may be probed by scattering at very high energies, the constituents of hadrons (quarks and gluons) behave almost as if they were free particles, originally called partons. Similarly, at larger distances scales α_{QCD} becomes larger, and perturbation theory breaks down. This increase in the coupling constant at large distances leads to a phenomenon known as confinement, which prevents colored objects (such as quarks or gluons) from being isolated. *See* GLUONS; QUANTUM CHROMODYNAMICS; QUARKS.

Gravity. Gravitation, as defined by the general theory of relativity formulated by A. Einstein in 1915, is also a gauge theory. In gravity the gauge transformations are the general coordinate transformations. Perturbatively, the gravitational interaction may be interpreted as the result of graviton exchange. The graviton is the spin-2 gauge particle associated with general coordinate invariance. Unfortunately, perturbative gravity is a nonrenormalizable theory, and it has not proved possible to formulate a consistent quantum theory of gravity. String theory shows some promise. By replacing point particles with strings, the divergences in higher orders of perturbation theory are softened. The fact that superstring theories are perturbatively renormalizable and have spin-2 excitations suggests that string theory might provide a correct theory of quantum gravity. However, at this point the understanding of string theory has not progressed far enough to determine whether this hope will be realized. To formulate the correct theory of quantum gravity remains one of the great challenges of theoretical physics. *See* GRAVITATION; QUANTUM GRAVITATION; RELATIVITY; SUPERSTRING THEORY.

Symmetry. One of the beautiful aspects of gauge theory is the overarching role that symmetry plays. Gauge symmetry is more than a simplifying principle that places restrictions on the possible dynamics of a theory; rather, it is the symmetry itself that determines the interactions. *See* ELEMENTARY PARTICLE; QUANTUM FIELD THEORY; SYMMETRY LAWS (PHYSICS).

Martin Bucher

Bibliography. E. Commins and P. Bucksbaum, *Weak Interactions of Leptons and Quarks*, 1983; C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, 1980; A. Pais, *Inward Bound*, 1986; C. Quigg, *Gauge Theories of Strong, Weak, and Electromagnetic Interactions*, 1983, reprint 1997; J. Schwinger (ed.), *Selected Papers on Quantum Electrodynamics*, 1958; F. Wilczek, Quantum chromodynamics: The modern theory of the strong interaction, *Annu. Rev. Nucl. Part. Sci.*, 32:177, 1982.

Gauss' theorem

The assertion, under certain light restrictions, that the volume integral through a volume V of the divergence of vector function $\Im(x, y, z)$ is equal to the surface integral of the exterior normal component of $\Im$ over the boundary surface S of V, or in symbols $\int\int\int \boldsymbol{\nabla} \cdot \Im \, dV = \int\int \Im \cdot \boldsymbol{\nu} \, dS$, with $\boldsymbol{\nu}$ the unit exterior normal to S. This theorem is also known as the divergence theorem and as Green's theorem.

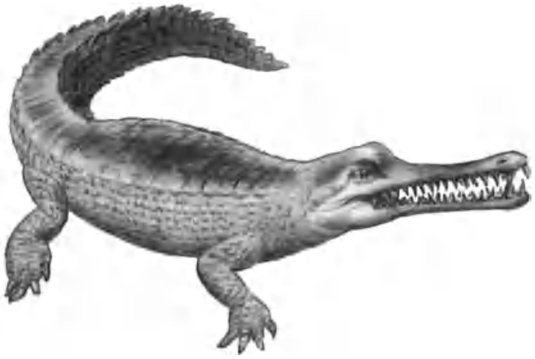
A simple illustration is provided by the radius vector $\mathfrak{r}$ ($\mathfrak{r} = xi + yi + zf$), and the sphere $S(O, a)$ (center at O, radius a). Because $\boldsymbol{\nabla} \cdot \mathfrak{r} = 3$ and on S, $\boldsymbol{\nu} = \mathfrak{r}a$, $\mathfrak{r} \cdot \boldsymbol{\nu} = a$, it follows that $\int\int\int \boldsymbol{\nabla} \cdot \mathfrak{r} \, dV = 3(4/\!\sqrt{3}\pi a^3)$, while $\int\int \mathfrak{r} \cdot \boldsymbol{\nu} \, dS = a(4\pi a^2)$.

A common method of proof consists of the conversion of $\int\int\int, \nabla \cdot \mathfrak{F}\, dV$ into a sum I of three double integrals with domains in the coordinate planes followed by the conversion of I to $\int\int \mathfrak{F} \cdot \boldsymbol{\nu}\, dS$. The function $\mathfrak{F}$ may be assumed to be of class C' and the volume V to be convex with smooth boundary pieces. The class of admissible volumes may be extended by the introduction of protuberances and cavities of the previous type.

The divergence theorem plays an important role in a variety of subjects such as electricity and magnetism, mechanics of continuous media (including fluid dynamics), heat flow, partial differential equations, and potential theory. *See* CALCULUS OF VECTORS; GREEN'S THEOREM. Homer V. Craig

Gavial

The name of two species of reptiles which form the family Gaviidae in the order Crocodilia. The Indian gavial (*Gavialis gangeticus*) is confined to the Ganges River and its tributaries and may attain a length of 30 ft (9 m), while the Malayan gavial (*Tomistoma schlegeli*) is found in Borneo and Sumatra and has a maximum length of about 15 ft (4.5 m).



Indian gavial (*Gavialis gangeticus*).

The gavial is distinguished from other members of the order by its extremely long, slender snout (see **illus.**). The tip of the snout is enlarged, and fleshy elevations surround the openings of the nostrils. The mouth cavity is lined with long, stout, sharp-edged teeth, the upper set interlocking with the lower. These animals feed almost exclusively on fish. They are timid animals and, when alarmed, move with great rapidity into the nearby water.

The females lay about 30–40 eggs in nests which they have prepared on the riverbanks. The newly hatched young are about 1 ft (0.3 m) long and quite active. *See* ALLIGATOR; CROCODILE; CROCODYLIA; REPTILIA. Charles B. Curtin

Gaviiformes

A small order of aquatic birds that contains a single living family, the Gaviidae (loons), with five species restricted to the Northern Hemisphere, and a fossil family, the Colymboididae. *See* AVES.

Common loon (*Gavia immer*), Jammer Lake, Minnesota. (© *DonGettyPhoto.com.*)

Phylogeny and fossil record. Relationships of the loons to other birds is uncertain; suggested affinities with the charadriiforms, the fossil hesperornithiforms, or the grebes are disputed. The earliest fossil loons may be *Polarornis* from the Cretaceous of Antarctica. Later forms are the small, less specialized *Colymboides*, known from the Eocene and Miocene of Europe and placed in a distinct family, the Colymboididae. Modern loons of the genus *Gavia* are known from the lower Miocene. *See* CHARADRIIFORMES; HESPERORNITHIFORMES; PODICIPEDIFORMES.

Characteristics. Loons, or divers, are large swimming and diving birds that catch fish with their strong bills in underwater pursuits. They are foot-propelled swimmers, with short legs placed far back on the body and webbed feet; land locomotion is awkward. Loons take wing with difficulty but are strong fliers. Sexes are similar in plumage, which is white below and checked black and white above (see **illustration**). Loons are monogamous, have a strong pair bond, and engage in elaborate courtship rituals that include loud, quavering calls often given in flight. Breeding is solitary and takes place on freshwater lakes. The young are downy, leave the nest immediately after hatching, and are cared for by both parents. Because all flight feathers are molted simultaneously, the birds are rendered temporarily flightless. Loons migrate south to wintering areas, which are mostly along ocean coasts.

Status. Common loons (*Gavia immer*; see illustration) in the southern part of their breeding range have been decreasing in number because of direct disturbance from boating on lakes and the effects of mercury poisoning as a consequence of chemical action of acid rain on deposits in lakes. *See* ENDANGERED SPECIES. Walter J. Bock

Bibliography. C. Carbonear, Order Gaviidae (divers), in J. del Hoyo et al. (eds.), *Handbook of* *the Birds of the World*, vol. 1, pp. 162–172, Lynx Edicions, 1992.

Gear

A machine element used to transmit motion between rotating shafts when the center distance of the shafts is not too large. Toothed gears provide a positive drive, maintaining exact velocity ratios between driving and driven shafts, a factor that may be lacking in the case of friction gearing which is subject to slippage. While the motion transmitted to mating gears is kinematically equivalent to that of rolling surfaces identical with the gear pitch surfaces, the action of one gear tooth on another is generally a combination of rolling and sliding motion. When the distance between shafts is large, other methods of transmission are used. *See* BELT DRIVE; CHAIN DRIVE; ROLLING CONTACT.

The application of gears for power transmission between shafts falls into three general categories: those with parallel shafts, those for shafts with intersecting axes, and those whose shafts are neither parallel nor intersecting but skew (see **table**).

Principal features. **Figures 1** and **2** illustrate the principal features of toothed gears. Such terms as pitch circle, addendum circle, and root circle, being geometrical, are defined by the diagrams.

Definitions for other commonly used terms for describing gears are addendum, the radial distance between the pitch circle and the addendum circle; dedendum, the radial distance between the pitch circle and the root circle; face, the tooth surface outside the pitch circle; flank, the tooth surface between the pitch circle and the root circle; clearance, the amount by which the dedendum exceeds the addendum of the mating gear; whole depth, addendum plus dedendum; and working depth, whole depth minus clearance; circular pitch, the distance from a point on one tooth to the corresponding point on the next, measured along the pitch circle; and base pitch, the similar measurement along the base circle.

Terms descriptive of a pair of spur gears in mesh are shown in Fig. 2. Where the pitch circles are tangent is pitch point *P*. Each pair of teeth always has its point of contact on the pressure line. The angle between this pressure line and the common tangent to the pitch circles through *P* is the pressure angle. As Fig. 2 illustrates, the pressure angle at which two gears actually operate depends on the distance between centers at which they are mounted. Initial contact between each pair of teeth occurs at point *C*, where the addendum circle of the driven gear crosses the pressure line; and final contact is at *D*, where the driver's addendum circle crosses the pressure line. Distance *CD* is the length of the path of contact. It must exceed the value of the base pitch if a pair of teeth is to come into contact before the previous pair has gone out of contact. The ratio of the length of the path of contact to the base pitch is called the contact ratio, which can be viewed as approximately the average number of pairs of teeth in contact. It is usual to design for a contact ratio of 1.4 or above to ensure smooth, continuous tooth action.

Chief types of application of gears for power transmission

Shaft relationship	Type of gearing
Parallel	Spur Helical Herringbone Double helical
Intersecting	Bevel
Skew	Helical Worm Hypoid

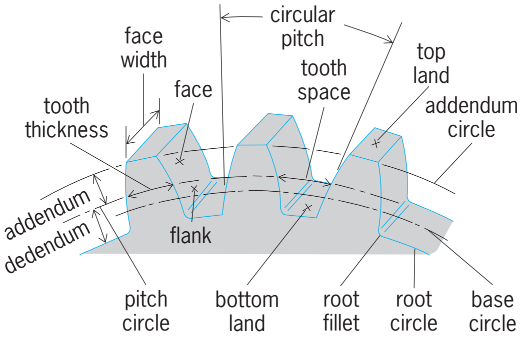


Fig. 1. Principal features of gear teeth.

Gear tooth sizes are designated by diametral pitch, which is the number of teeth per inch of diameter of the pitch circle. Pitch circle is, in turn, the circle whose periphery is the pitch surface, or surface of an imaginary cylinder that would transmit by rolling contact the same motion as the toothed gear. A gear with a 20-in. (51-cm) pitch diameter and teeth having a diametral pitch of 2 has 40 teeth. Diametral pitch *P* times circular pitch *p* (Fig.1) equals π. The diametral pitches most widely used are: 1, 11/2, 2, 21/2, 3, 4, 5, 6, 8, 10, 12, 16, 20, 24, 32, 40, 48, 64, 72, 96, 120, 160.

The smaller of two gears in mesh is the pinion. It has the fewer teeth and is the driving gear in a speed reduction unit. The minimum number of teeth an involute pinion can have and still run without interference between its flanks and the tips of the mating gear teeth is fixed by the tooth system. Smaller pinions are possible only if the pinion's flanks are undercut.

Backlash. The amount of which the tooth space of a gear exceeds the tooth thickness of the mating gear at the pitch circle is the backlash. It can be determined in the plane of rotation or, for helical gears, in the plane normal to the tooth face.

If mating gears have zero backlash, gears and mountings need to be dimensionally perfect. To retain zero backlash with varying operation conditions, all parts need exactly the same thermal expansion characteristic. Because of the difficulty of meeting these requirements and for lubrication, freedom-backlash is provided between gear teeth. The usual practice is to reduce the tooth thickness of each gear by an amount equal to half the desired backlash. However, in the case of a gear and a small pinion it is customary to reduce the tooth thickness on the gear only, leaving the pinion with standard tooth thickness. Backlash can also be adjusted by slight changes in the center distance between gears. Except for a small change in the pressure angle, the action of involute gear teeth is not affected by backlash or center distance adjustment.

In the case of precision gearing for control systems and similar applications, backlash results in a nonlinear relation between input and output. Several methods of reducing backlash are in use. One method is to place two identical spur gears on the same shaft, one fixed to the shaft, one free. The loose gear is attached to the fixed one by springs, which keep the composite gear in positive contact with the pinion at all times. A second method is to use tapered-tooth gearing (beveloid) with adjustment along the shaft to eliminate excessive backlash.

Gear action. A principal function of gears is to change the speed of rotation. This action is described by the velocity ratio of the gears in mesh. Ratio VR is the number of revolutions N_1 of the driving gear divided by the number of revolutions N_2 of the driven gear in the same time interval. For gears with teeth T_1 and T_2, respectively, VR is expressed as in Eq. (1).

$$VR = \frac{N_1}{N_2} = \frac{T_2}{T_1} \qquad (1)$$

When two curved surfaces, such as the mating surface of two gear teeth, are in driving contact, there is a definite velocity ratio between the bodies. The angular velocities of the two bodies are inversely proportional to the segments into which their line of centers is divided by a line passing through their point of contact and normal to their surfaces at this point. Thus a constant angular velocity ratio between bodies in driving contact demands that the common normal to the profiles at the point of contact cut the center line at a fixed point, the pitch point. This latter statement is frequently referred to as the fundamental law of gear tooth action. Pure rolling contact between gear teeth occurs only when they are in contact at the pitch point. At all other positions the teeth have some sliding with the maximum sliding at the first and last instants of contact. Although there are a number of tooth forms that will satisfy the fundamental law, only two of them have been used to any great extent. The cycloidal tooth predominated until the late 1800s but has been replaced to a great extent by the involute gear tooth. Cycloidal teeth are still found in instruments, watches, clocks, and, occasionally, in cast and cut gears.

Gears are said to be interchangeable when any gear of the set will run with any other gear of the same set. Actually there is little need today for gears that are interchangeable. In the manufacture of the overwhelming proportion of machines pairs or groups of gears are designed to mesh with each other and no others. Standardization in manufacture is common, however. Each of the gears is made from one of a set of standardized cutters. Thus it is the standardization of the production tools rather than of the gears themselves that is of most importance.

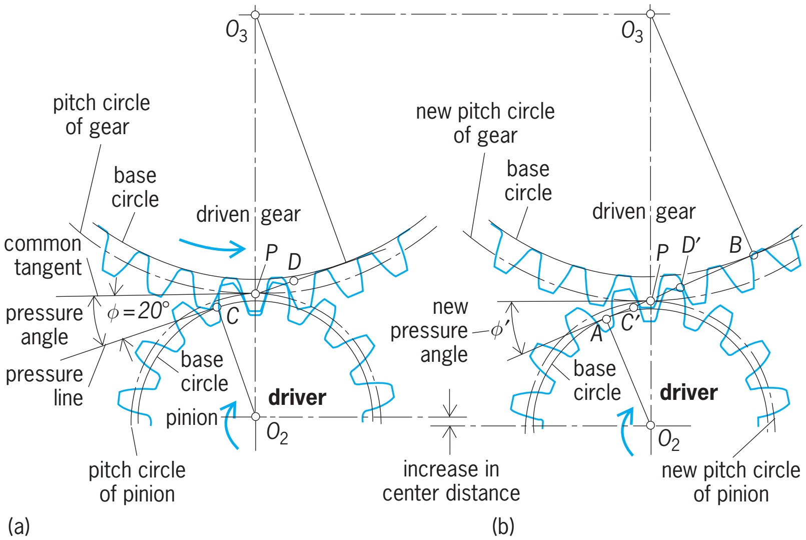


Fig. 2. Action of involute gearing. (*a*) Gear and pinion mounted at normal center distance. (*b*) Gear and pinion mounted at greater than normal center distance. (*After J. E. Shigley, Theory of Machines, McGraw-Hill, 1961*)

Involute gear teeth. An involute tooth is laid out along an involute (**Fig. 3**), which is the curve generated by a point on a taut wire as it unwinds from a cylinder. The generating circle is called the base circle of the involute. Proportions of the tooth are fixed by the gear tooth system and the diametral pitch. The involute curve establishes the tooth profile outward from the base circle. From the base circle inward, the tooth flank ordinarily follows a radial line and is faired into the bottom land with a fillet. The basic rack form of the involute tooth has straight sides, an important property from the manufacturing standpoint.

Gear teeth may interfere with each other, especially where pinions have a relatively small number of teeth. In Fig. 2, point of initial contact C occurs on the pressure line to the right of (after) the point of tangency of the pressure line and the pinion's base

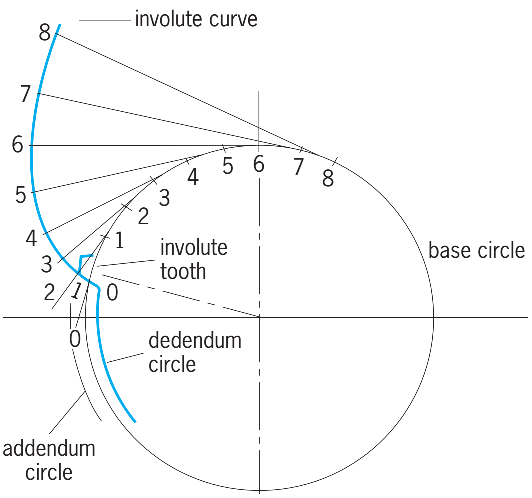


Fig. 3. Method of generation for face of involute tooth.

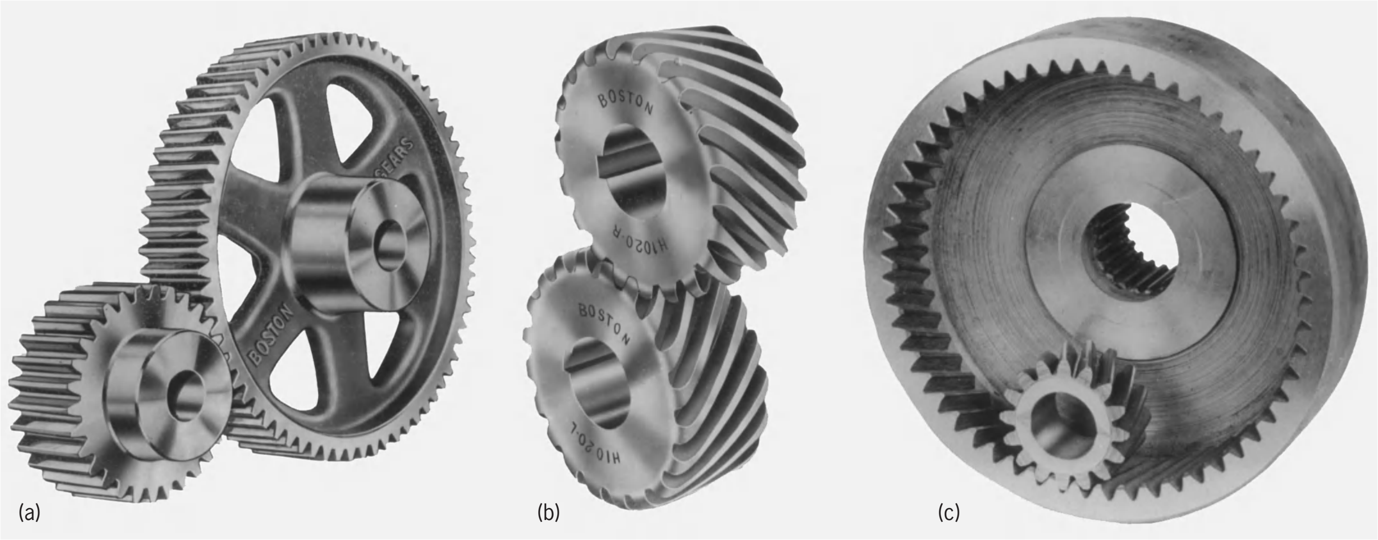


Fig. 4. Spur gears. (*a*) External spur gear and pinion, the commonest type of spur gear. (*b*) Single helical gearset (*Boston Gear Works*). (*c*) Helical internal gear and pinion (*Fellows Gear Shaper Co.*).

circle. Here there is no interference. But if point *C* were to the left of the tangency point, this would indicate premature contact between the teeth, a contact occurring on the noninvolute portion of the pinion tooth flank below the base circle. The tip of the gear tooth, in such a case, digs into the flank of the pinion tooth.

There are several ways to eliminate interference: (1) Reduce addendum of the gear. (2) Increase pressure angle. (3) Increase backlash by increasing center distance between gears. (4) Undercut flank of the pin ion. (5) Relieve or modify face of the gear tooth.

Spur gears. In the truest sense, spur gears are only those that transmit power between parallel shafts and have straight teeth parallel to the gear axis (**Fig. 4***a*). It is common practice, however, to group helical gears that have parallel shafts under the heading of spur gears (Fig. 4*b*). The pitch surfaces of gears with parallel axes are rolling cylinders, and the motion these gears transmit is kinematically equivalent to that of the rolling pitch cylinders. Spur gears are classified as external, internal, and rack and pinion. External spur gears, the most common, have teeth which point outward from the center of the gear. Internal or annular gears have teeth pointing inward toward the gear axis (Fig. 4*c*). A rack may be considered as a gear having an infinite pitch circle radius. Thus its pitch surface is a plane. A rack and pinion running together transform rectilinear motion into rotary motion, or vice versa.

To standardize manufacture of gears, the American Gear Manufacturers Association has adopted three basic types of spur-gear tooth, with pressure angles of 20° and 25° for coarse-pitch gears (with a diametral pitch under 20) and 20° for fine-pitch gears (diametral pitch of 20 and above).

An internal gear has the positions of the addendum and dedendum reversed from those of an external gear. This results in a different tooth action and less slippage than with an equivalent external spur gear. For a given tooth ratio, the arc of action of an internal gear is slightly greater than that of an external gear of the same size and the tooth is stronger. The nature of an internal gear makes it especially suited to closer center distances than could be used with

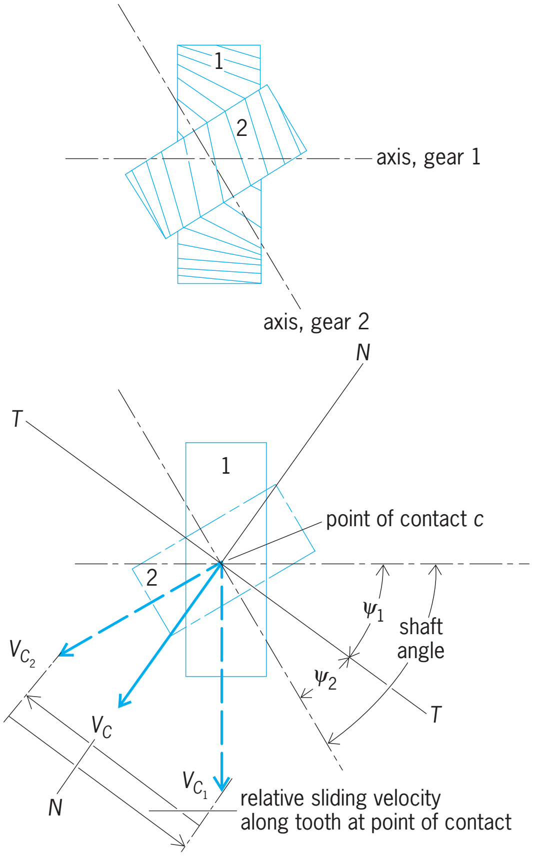


Fig. 5. Action of crossed helical gears.

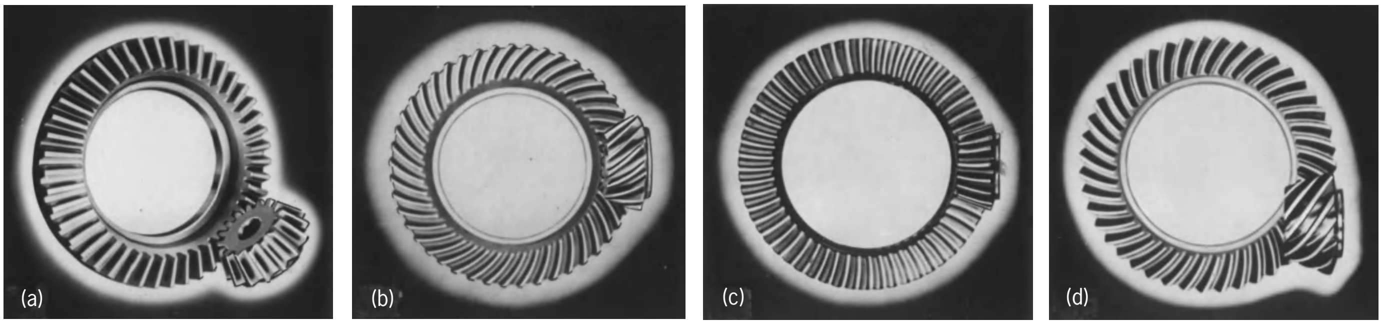


Fig. 6. Bevel gears with pinions. (*a*) Straight. (*b*) Spiral. (c) Zero. (*d*) Hypoid. (*Gleason Gear Works*)

an external gear of the same size. When it is necessary to maintain the same sense of rotation for two parallel shafts, the internal gear is especially desirable because it eliminates the need for an idler gear. These conditions make the internal gear highly adaptable to epicyclic and planetary gear trains.

Noncircular gears are used to obtain velocity ratios that vary in a precise manner. Elliptical gears are an example of noncircular gears. They provide a convenient method of obtaining a quick return for machines that do most of their work during only a portion of the drive shaft revolution. Noncircular gears are used in computing mechanisms and other devices where a prescribed varying output function is to be obtained using a linear input.

Helical gears. Gears running on parallel axes and with teeth twisted oblique to the gear axis are essentially spur gears. Because of the twist, contact is progressive across the tooth surface, starting at one edge and proceeding across the face of the tooth. The action results in reduced impact and quieter operation, particularly at high speed. Herringbone gears are equivalent to two helical gears of opposite hand placed side by side. They are especially suited for high-speed operation and eliminate the axial thrust produced by single helical gears. Double helical gears have a central groove for tool runout, making it possible to finish the teeth by a shaving operation. They can be run at even higher speeds than herringbone gears. Dimensions of a helical gear are determined on both the plane of rotation and on the plane perpendicular to the helix angle of the tooth. Thus, a helical gear has a circular pitch measured on the plane of rotation and a normal circular pitch measured on the normal plane.

Crossed helical gears. Where shafts cross obliquely, motion is transmitted by crossed helical gears (**Fig. 5**). The teeth are helical but differ from the teeth of worm gears in that no one tooth (thread) makes a complete turn on the pitch circle. Pitch surfaces of crossed helical gears are cylindrical as with spur gears. However, with crossed shafts, the oblique teeth have point contact rather than the line contact that occurs with parallel shafts. Analysis of the gears is based on equal components of the pitchpoint velocity on each mating gear along common normal N–N. Sliding occurs in direction T–T of the tooth elements. As with spur gears, the revolutions per unit time are inversely proportional to the numbers of teeth. Velocity ratio VR may also be expressed as in Eq. (2), where D is pitch diameter and ψ helix angle.

$$VR = \frac{N_1}{N_2} = \frac{D_1 \cos \psi_1}{D_2 \cos \psi_2} \tag{2}$$

Helix angle is between the shaft axis and a line tangent to the tooth through the pitch point.

Helical gears are referred to as right- or left-hand in the same manner as screw threads, a right-hand gear being one on which the teeth twist clockwise as they recede from an observer looking along the axis.

Bevel gears. Where shafts intersect, bevel gears transmit the motion. Such gears may be used only to change the shaft axis direction or to change speed as well as direction. Two bevel gears with equal numbers of teeth and running together with their shaft axes intersecting at 90° are called miter gears. Several forms of bevel gears are in use, including straight-tooth, spiral, and skewed bevel gears (**Fig. 6**).

External bevel gears have pitch angles less than 90° (**Fig. 7*a***). Internal bevel gears have pitch angles greater than 90°, hence their pitch cones are inverted (Fig. 7*b*). A crown gear is one having a pitch angle of 90° (Fig. 7*c*). Thus its pitch surface is a plane, and the crown gear corresponds in this respect to a rack in spur gearing.

Straight bevel gears. The simplest form of bevel gear has straight teeth which, if extended inward, would come together at the intersection of the shaft axes. This point is also the apex of the rolling cone, which forms the pitch surface of the gear. Much of the terminology applied to bevel gears is the same as that used for spur gears. Additional terms used in reference to bevel gears are given in **Fig. 8**. Diametral

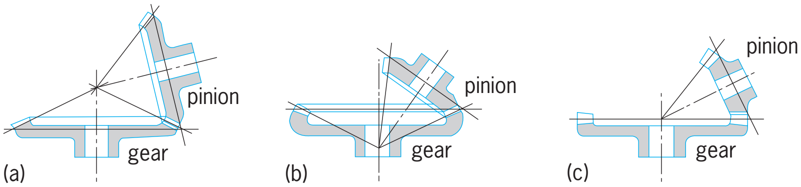


Fig. 7. Bevel gears. (*a*) External. (*b*) Internal. (c) Crown. (*After G. L. Guillet and A. H. Church*, *Kinematics of Machines*, *5th ed.*, *Wiley*, *1950*)

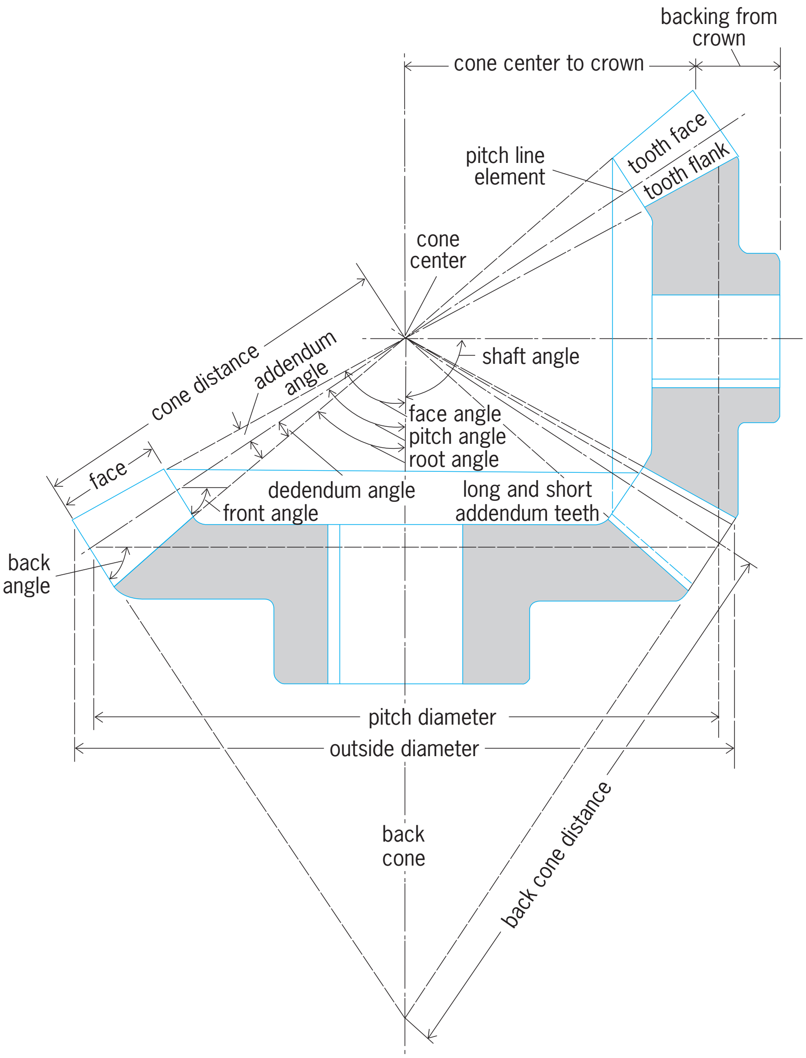


Fig. 8. Characteristics of bevel gears. (*After C. W. Ham, E. J. Crane, and W. L. Rogers, Mechanics of Machinery, 4th ed., McGraw-Hill, 1948*)

pitch of a bevel gear is not constant across the full width of the tooth. The diametral pitch at the pitch diameter is used in fixing tooth proportions. The formative number of teeth in a bevel gear is the number of teeth that would be a spur gear whose pitch radius equaled the back cone distance of the bevel gear.

Speeds of the shafts of bevel gears (velocity ratio) are inversely proportional to the numbers of teeth on the gears or to the sines of the pitch angles, but not to the formative number of teeth. Use of straight bevel gears is limited to low-speed operations, ordinarily below 1000 surface feet per minute or, in the case of small gears, 1000 rpm.

Because each point on a straight tooth bevel gear remains a fixed distance from the pitch cone apex, there is no sliding along the tooth as it engages. Contact across the full tooth face occurs instantaneously—as with spur gears—as the teeth come into mesh.

Spiral bevel gears. To provide a gradual engagement, as contrasted to the full line engagement of straight bevel gears, the teeth of spiral bevel gears are curved and oblique. Theoretically the curve is a spiral but, to facilitate manufacture, the curve is actually a circular arc which, within the tooth face width, closely approximates a spiral. This tooth inclination brings more teeth in contact at any one time than with an equivalent straight-tooth bevel gear. The result is smoother and quieter operation, particularly at high speeds, and greater load-carrying ability than with straight bevel gears of the same size.

Spiral bevel gears are used in sewing machines, motion picture equipment, machine tools, and other applications where quiet, smooth operation is essential. They should, in general, be mounted on antifriction bearings because of the axial thrust due to the oblique teeth. In the past they have been used extensively in the rear axle drives of automobiles, but are being replaced by hypoid gears.

Zero bevel gears. A special form of bevel gear has curved teeth with a zero-degree spiral angle. Thus the teeth are not oblique as is the case with spiral bevel gears. Rather, the teeth lie in the same general direction as those of an equivalent straight-tooth bevel gear, and so the gears are usually used in the same types of drives as the straight-tooth gear. As with straight bevels, they produce no axial thrust and, therefore, may be used without thrust bearings. The face that they may be produced on the same equipment as spiral, bevel, and hypoid gears makes them economically desirable.

Tooth proportions for bevel gears follow the standards established by the Gleason Works and adopted as the recommended standard of the American Gear Manufacturers Association. Tooth proportions are a function of the velocity ratio. Thus, bevel gears are not as interchangeable as spur gears.

Hypoid gears. To connect nonparallel, nonintersecting shafts, usually at right angles, hypoid gears are used. They are similar to spiral bevel gears in their general appearance. The axis of the hypoid pinion may be offset above or below the axis of the gear. The shape of the tooth is similar to that of the spiral bevel gear and gives progressive contact across the tooth. In operations these gears run even more smoothly and quietly than spiral bevel gears. To maintain line contact of the teeth, with the offset shaft, the pitch surface of the hypoid gear is a hyperboloid of revolution rather than a cone as in bevel gears.

One of the first uses of hypoid gears was in the rear-axle drive of Packard automobiles. The operating smoothness of hypoid gears, along with the lower body lines made possible by the offset pinion shaft, has made them extremely popular for automotive use. Industrial applications of the hypoid gear also take advantage of the pinion offset, which allows the mounting of any number of pinions on a single continuous shaft, a feat not possible with bevel gears. The shaft arrangement of the hypoid gear and pinion enables bearings to be placed on both sides of the gear and of the pinion. The offset axis results in a larger and, consequently, stronger hypoid pinion tooth than on an equivalent straight-tooth or spiral bevel gear. Expressed in terms of pitch diameter, a hypoid pinion has fewer teeth than a spiral bevel pinion of the same pitch diameter. It is possible to use hypoid pinions having 7, 8, or 9 teeth in contrast to a minimum of 12, 13, or 14 teeth—depending on the velocity ratio—on a spiral bevel pinion. Lubricants must withstand the higher loading and the sliding that occurs along the teeth of hypoid gears.

Hypoid gears are suitable for large velocity reductions; reduction ratios of 60:1 and higher are entirely feasible. In general, shaft offset should not exceed 40% of the equivalent bevel gear back cone distance and, when the loading is heavy, as in truck and tractor drives, the offset should be nearer 20% of this distance. Direction of the offset, above or below center, must be specified for any given installation. The gears should, in general, be mounted on antifriction bearings in an oil-tight case. Thrust bearings must be provided. Because of the sliding tooth action, the efficiency of hypoid gears is somewhat less than that of equivalent bevel gears.

Worm gears. A chief way to connect nonparallel, nonintersecting shafts that are at right angles is through a worm gear. The worm, ordinarily the driver, is similar to a crossed helical gear except that it has at least one complete tooth (thread) around the pitch surface. The mating gear is the worm wheel or worm gear. Worm gearing is generally used to obtain large velocity reductions with the worm as the driver and the worm wheel as the driven gear, although occasional applications, for example cream separators, have the worm wheel as the driver. The pitch surfaces of straight worms are cylinders and the involute teeth have point contact. Because the appearance of the worm is similar to that of a screw, the teeth are called threads.

The pitch of the worm is the axial distance from any point on one tooth to the corresponding point on the next tooth (**Fig. 9**). This must equal the circular pitch of the mating worm wheel. Lead is the axial distance the worm helix advances in one complete revolution around the pitch surface. A single thread worm has the pitch and lead equal; one revolution of such a worm will, for shafts at right angles, advance the worm wheel $1/N$ revolutions if N is the number of teeth on the worm wheel. A double-threaded worm has the lead equal to twice the pitch and will then advance the worm wheel $2/N$ revolutions per turn of the worm. Thus, worm gears follow the general rule of angular velocity ratio inversely proportional to the ratio of the numbers of teeth. Worms are right- and left-hand in the same sense as helical gears. Changing hand of the worm reverses the relative rotation of the worm wheel.

Improved load-carrying capacity and wear characteristics are made possible by increasing the contact between the worm and wheel (**Fig. 10**). Line contact is obtained by making the worm wheel surface concave to conform to the tooth profile of the worm. Still greater contact is obtained by using a concave worm as well. Known as a cone-drive or Hindley worm, this design permits greater contact surface and allows more teeth to be in contact at one time. *See* GEAR TRAIN; MECHANISM; PLANETARY GEAR TRAIN.

John R. Zimmerman

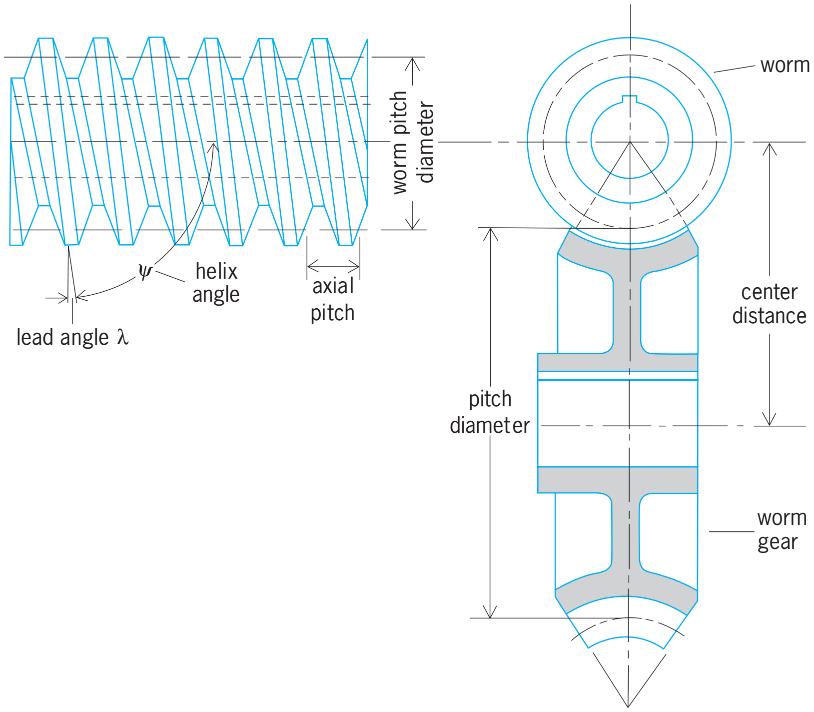


Fig. 9. Nomenclature of a single-enveloping worm gearset. Teeth of worm gear are called threads. (*After J. E. Shigley*, *Theory of Machines*, *McGraw-Hill*, *1961*)

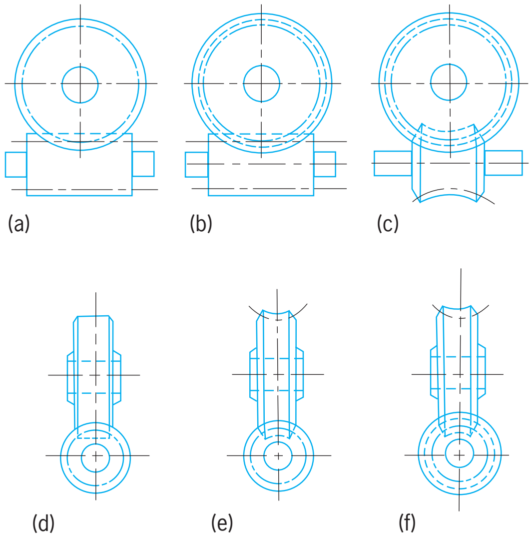


Fig. 10. Nature of contact for worm gears. (*a*) Nonthroated. (*b*) Single-throated. (*c*) Double-throated on cone. (*d*) Point contact. (*e*) Line contact. (*f*) Area contact. (*Michigan Tool Co.*)

Bibliography. N. P. Chironis (ed.), *Gear Design and Application*, 1967; N. P. Chironis, *Mechanisms and Mechanical Devices Sourcebook*, 3d ed., 2001; J. R. Colbourne, *The Geometry of Involute Gears*, 1987; J. W. Patton, *Kinematics*, 1979; B. A. Shtipelman, *Design and Manufacture of Hypoid Gears*, 1978; D. W. South, *Encyclopedic Dictionary of Gears and Gearing*, 1994.

Gear loading

The power transmitted or the contact force per unit length of a gear. If gear speed, size, and tooth contour are fixed, the power-handling capacity may be increased by increasing the axial length of the gear. However, where space is restricted (as in aircraft and many ground vehicles), the gear is usually loaded to its safe limit.

To obtain satisfactory tooth-surface durability from highly loaded gears, experience indicates that several items of the gear set must be properly designed and manufactured, namely, the tooth profile which must be properly modified from a true involute to suit the operating conditions; index of teeth and parallelism of teeth, which must be held within close limits; gearing, which must be mounted so the teeth will not deflect out of line; and gear tooth surfaces, which must be of sufficient hardness and proper finish and which should have good lubrication, particularly on start of initial operation.

Gear design is a compromise between tooth strength and surface durability. Larger teeth give more tooth strength but less surface durability; smaller teeth result in less tooth strength but more surface durability.

Surface distress on the teeth becomes particularly important at loadings in the order of thousands of pounds per inch of face.

In designing the gear tooth surfaces, one mesh of teeth is assumed to carry the load through the action, and the action is balanced so that the product of Hertz maximum compressive pressure P in pounds per square inch and sliding velocity V in feet per second is the same at each end of the assumed single-mesh action. This product, known as the PV value, may be 3,000,000 or higher in gears with high pitch-line speeds for which the designer may use PVT values. The design may produce a smaller contact ratio than is conventional, but this result indicates that contact ratio alone is an unreliable measure of gear capacity.

A highly loaded tooth of adequate rigidity deflects about a point in the middle of the rim, bending as a rigid body under load rather than as a nonuniform beam only. Relief or other modification of tooth profile provides clearance so as to avoid excessive loading at teeth tips due to deflection of the preceding mesh, and ramps at the tooth tips assure that first contact does not extend to the tips.

This design refinement necessitates corresponding care in production, as in minimizing distortion during carburizing. Teeth can be held parallel within 0.0003 in. (7.5 micrometers) in the width of the tooth, and the index can be maintained within 0.0002 in. (5 μm) between adjacent teeth of a gear.

To assure removal of all possible error in grinding when this method of finishing is used and to achieve a more satisfactory working surface on the gear tooth face, grinding wheels are dressed to produce a finish which is coarser than usual. Slight surface roughness (profilimeter readings between 15 and 37) produces less surface distress, or scuffing, of highly loaded steel gears than do smoother finishes, although a steel gear running with a bronze one should be smoother.

Quiet operation is an indication of efficient operation free from abrupt tooth engagement and severe scuffing of one tooth face by another. Helical gears provide smooth transition of effort from tooth to tooth. The gear is designed with a modified involute to provide soft contact at the start of engagement and end of engagement to reduce the pressure on the teeth when the sliding of the tooth surfaces is highest. For quietest operation they should be cut and shaved accurately for lead, profile, and index. A helical contact ratio of 2 or more is desirable. This ratio can be obtained by increasing the width or the helix angle, although effort expended to increase helical contact ratio reduces deflection and PV values more than does the same effort expended to increase involute contact ratio. As tooth size and resultant sliding increase, this proportioning of effort becomes more important. Helix angles in the order of 45° on constant mesh gears require more expensive mountings but improve quietness. Rigid and accurate mounting is essential to achieve best operating condition.

With other gears subject to deflection, such as hypoid rear-axle gearing, conventional practice is to crown the tooth and modify the tooth profile to keep contact away from the tooth ends and tips. The gear set is usually designed for stiffness rather than for strength, and the shafts and mountings should also be designed on this criterion so that little crowning is needed. *See* GEAR.

Forest R. McFarland

Bibliography. E. Buckingham, *Analytical Mechanics of Gears*, 1988; D. L. Landau, E. M. Lifshitz, and A. M. Kosevich, *Theory of Elasticity*, 3d ed., 1986, revised 1995; A. Stokes, *Gear Handbook*, 1992; J. P. Smith, *Gears and Their Vibration: A Basic Approach to Understanding Gear Noise*, 1983; D. Townsend, *Dudley's Gear Handboook*: *Design, Manufacture, and Application*, 1992.

Gear train

A combination of two or more gears used to transmit motion between two rotating shafts or between a shaft and a slide. In theory two gears can provide any speed ratio in connecting shafts at any center distance, but it is often not practical to use only two gears. If the ratio is large or if the center distance is relatively great, the larger of the two gears may be excessively large. Moreover, an additional gear may be necessary simply to give the proper direction to the output gear. Belt, rope, and chain drives are frequently used in conjunction with gear trains. *See* BELT DRIVE; CHAIN DRIVE; PLANETARY GEAR TRAIN.

Classification. Gear train classifications include simple, compound, reverted, epicyclic (planetary), and various combinations. The most important distinction is that between ordinary and epicyclic gear trains. In ordinary trains (**Fig. 1***a*), all axes remain stationary relative to the frame. But in epicyclic trains (Fig. 1*b*), at least one axis moves relative to the frame.

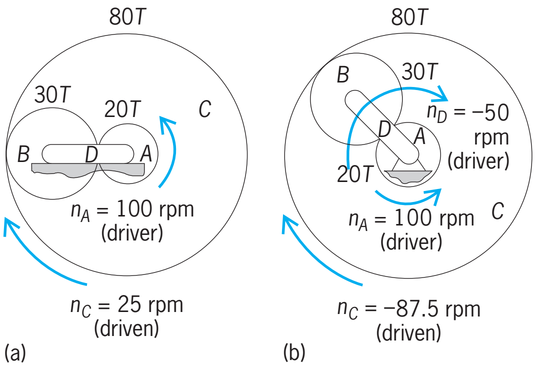
In Fig. 1*b* gear *B*, whose axis is in motion, is called a planet. The gears *A* and *C* are sun gears.

An ordinary gear train is a single degree of freedom mechanism: A single input, such as an input to gear *A* of the train in Fig. 1*a*, suffices to control the motions of the other moving members. But an epicyclic gear train (Fig. 1*b*) has two degrees of freedom: Two inputs are necessary. In the epicyclic train of Fig. 1*b*, the two input members are gear *A* and the planet carrier, link *D*. Only if both these members are controlled by external agencies can the motions of gears *B* and *C* be predicted. Frequently one gear of an epicyclic train is fixed. This then is one of the input members with a velocity of zero revolutions per unit time.

A simple gear train is one in which each gear is fastened to a separate shaft (Fig. 1*a*). If at least one shaft has two or more gears fastened to it (**Fig. 2**), the train is compound. The train of Fig. 2 is also a reverted gear train, because the input and output shafts are in line. If the input shaft is not in line with the output shaft, the train is nonreverted.

The 1000-hp (750,000-W) mill drive and pinion stand in **Fig. 3** are an example of a large industrial ordinary gear train. The first speed reduction is with the opposed single helical gears, and the second is through the herringbone gears. This train is compound and nonreverted.

Ratios. For an ordinary gear train, the ratio of the angular velocity of the last driven gear to that of the driving gear is known as the train value. The ratio of the driver's velocity to that of the last driven gear is the velocity ratio. By these definitions train value and velocity ratio are reciprocal quantities. The train value for the ordinary train of Fig. 1*a* is −0.25. The output gear *C* turns at one-fourth the speed of the input gear *A* and in the opposite direction.



(a) (b)

Fig. 1. Gear trains. (*a*) Ordinary. (*b*) Epicyclic.

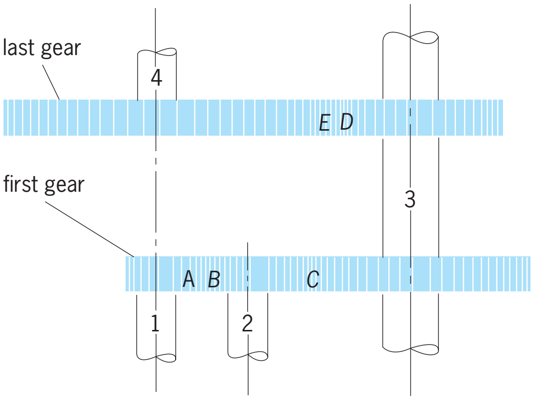


Fig. 2. Compound reverted gear train.

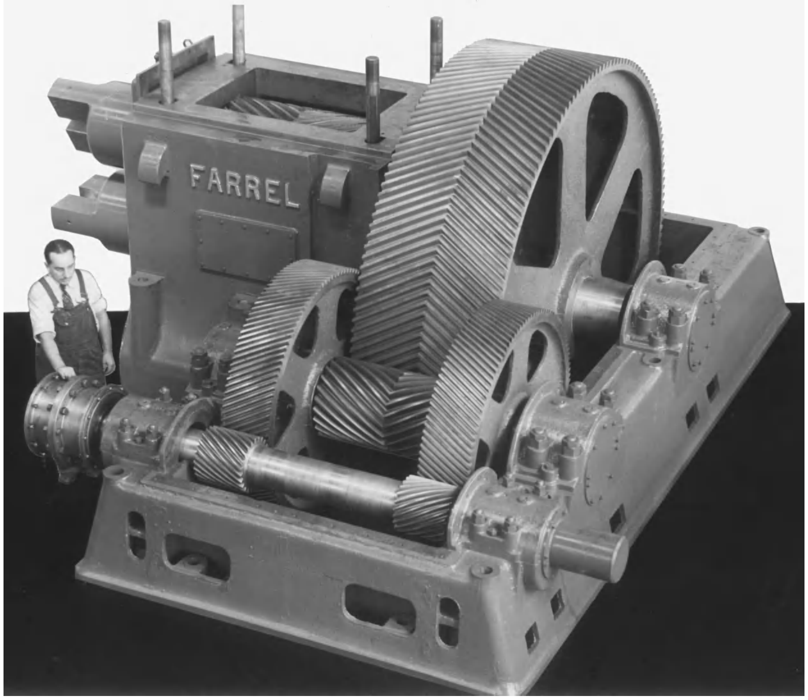


Fig. 3. Helical gear speed reducer. (*Farrel-Birmingham Co.*)

For epicyclic trains the situation can be a little more complicated. Since in general the output velocity is dependent on the two input velocities, the term train value is nonspecific. However, if one gear of an epicyclic train is fixed, the output velocity is some multiple of the velocity of the nonfixed input member; the term train value then applies.

Direction of rotation. The direction of rotation of any gear in an ordinary gear train can be determined by inspection. Mating external gears have opposite directions of rotation, while an internal gear has the same direction as its mating gear.

Ordinary train. Train value is by definition Eq. (1)

$$E = \frac{\text{angular velocity of the last driven gear}}{\text{angular velocity of the first driving gear}} \tag{1}$$

when the angular velocities are measured with respect to the frame supporting the gears. Referring to the ordinary train of Fig. 2, the train value is (ignoring sign) Eq. (2), where n is the speed in revolutions

$$E = \frac{\text{rpm of shaft 4}}{\text{rpm of shaft 1}} = \frac{n_4}{n_1} \tag{2}$$

per unit time. This can be expanded as an identity for the entire train to Eq. (3). Each of the ratios on

$$E = \frac{n_4}{n_1} = \frac{n_2}{n_1} \times \frac{n_3}{n_2} \times \frac{n_4}{n_3} \tag{3}$$

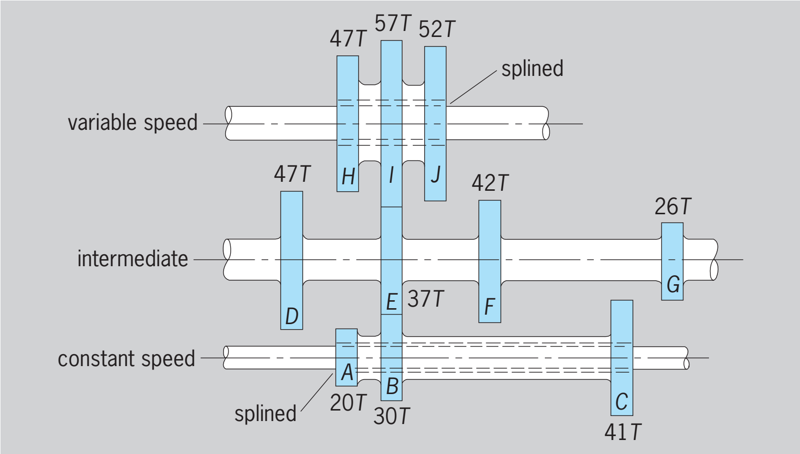


Fig. 4. Compact 9-speed transmission utilizing 10 gears and 3 shafts. (*After J. R. Zimmerman, Elementary Kinematics of Mechanisms, Wiley, 1962*)

the right side of Eq. (3) is the train value for a pair of meshing gears. Since for any two gears in mesh the speeds vary inversely as the numbers of teeth, the train value for the whole train must be (again ignoring sign) Eq. (4), where N is the number of teeth. In Eq. (4) the number of teeth appearing in the numerator of the expression on the right are those for driving gears, while those appearing in the denominator are the numbers of teeth on the driven gears. Thus, the general expression for the magnitude of the train value for an ordinary gear train is Eq. (5).

$$E = \frac{n_4}{n_1} = \frac{N_A}{N_B} \times \frac{N_B}{N_C} \times \frac{N_D}{N_E} \tag{4}$$

$$E = \frac{\begin{array}{c}\text{product of numbers}\\ \text{of teeth on driving gears}\end{array}}{\begin{array}{c}\text{product of numbers}\\ \text{of teeth on driven gears}\end{array}} \tag{5}$$

Because any two gears in mesh must have the same diametral pitch, expressions for the train value can also be written in terms of pitch diameters instead of numbers of teeth.

In the expression for the train value of the train of Fig. 2 the number of teeth on gear B cancels out. Gear B is an idler; it is both a driver and a driven gear. Its size has no effect on the train value's magnitude. It does affect, however, the sign of the train value. Idlers also are useful where a relatively large center distance must be spanned.

Transmissions. If a machine such as a machine tool or motor vehicle must be operated at any one of several output speeds, a multiple-speed gearbox, or transmission, may be used as a component part. The speed of the output shaft of a transmission can be varied by sliding gears in and out of contact or by connecting gears in continual mesh to shafts by means of clutches. A compact nine-speed transmission utilizing 10 gears and three shafts is shown in **Fig. 4**. Gears A, B, and C slide as a unit on the constant-speed shaft, providing three ways in which the constant speed and the intermediate shafts can be connected (A and D, or B and E, or C and G). Similarly gears H, I, and J slide as a unit on the variable-speed shaft, providing three ways in which it can be connected to the intermediate shaft (D and H, or E and I, or F and J). Consequently nine possible combinations are available. For example, the lowest output speed is attained when A and D mesh and E and I mesh. The train value is then as in Eq. (6).

$$E = \frac{20 \times 37}{47 \times 57} = 0.276 \tag{6}$$

The highest speed occurs when C and G mesh and D and H mesh. The train value for this combination is shown in Eq. (7).

$$E = \frac{41 \times 47}{26 \times 47} = 1.577 \tag{7}$$

See AUTOMOTIVE TRANSMISSION.

Epicyclic train. An epicyclic train is named for the path described by any point on the pitch circle of a planet gear as it rolls on a sun gear. The term epicyclic refers only to the motion and is not related to the gear tooth form, which may be involute or any other form satisfying the law of gearing. An epicyclic train can be used to obtain a considerably greater velocity ratio than would be possible with an ordinary gear train of the same size.

In the compound and reverted epicyclic gear train of **Fig. 5**, gear A is fixed to the frame, but not to shaft 1. The planet carrier, arm or link E, is fastened to shaft 1 and carries planet gears B and C, which are fastened to a common shaft. Planet B rolls on sun gear A, while planet C rolls on sun gear D. Gear A is one of the input members; it has a velocity of 0 rpm. The other input could be either the carrier E or gear D. In the example to follow it will be assumed that the arm carrier E is the input member.

To determine the train value of an epicyclic train with a fixed gear, the following special but very convenient technique can be used. The net motion is divided into two parts. In the first, the train is locked with no relative motion between any of its components. For this step gear A is released from the frame and locked to gear B. The locked train is rotated one revolution in an arbitrary direction. Counterclockwise rotation is assumed positive, clockwise negative. In the second step, the assembly is viewed as an ordinary gear train with the planet carrier regarded as fixed. While the fixed gear is returned to its original position by one revolution in the sense opposite that used in the first step, the consequent rotation

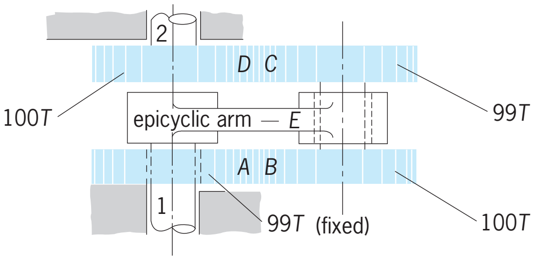


Fig. 5. Reverted epicyclic gear train.

Algebraic sum of revolutions for a reverted epicyclic gear train

Category	Gear A	Gear B	Gear C	Gear D	Arm E
Locked train	+1	+1	+1	+1	+1
Locked arm	−1	$+\frac{99}{100}$	$+\frac{99}{100}$	$+\frac{99}{100} \times \frac{99}{100}$	0
Net motion		$\frac{199}{100}$	$\frac{199}{100}$	$\frac{199}{10{,}000}$	+1

of the other gears is noted. For any one gear, the algebraic sum of the revolutions during the two parts is the net motion of that gear for one revolution of the planet carrier. To illustrate, consider the **table** for the train of Fig. 5, with the number of teeth as marked for each gear. The table shows that, for each revolution of the carrier, gear *D* will turn 199/10,000 of a revolution in the same direction.

For epicyclic trains with no fixed gears, a modification of the two-step method is needed. In the first step, the locked train is given the same rotation as the known motion of the planet carrier. In the second, with the carrier stationary, the gear whose velocity is known is rotated sufficient turns in the proper direction to make its net motion equal to the known value.

Probably the most common epicyclic train with no fixed gear is the automobile differential. When the automobile is moving along a straight path, there is no relative motion between the bevel differential gears fastened individually to the right and left axles. But as the car makes a turn, the differential gears move relative to one another in the manner of an epicyclic bevel gear train with no fixed gears. *See* DIFFERENTIAL; GEAR; PLANETARY GEAR TRAIN. John R. Zimmerman; Donald L. Anglin

Bibliography. H. H. Mabie and C. F. Reinholtz, *Mechanisms and Dynamics of Machinery,* 4th ed., 1987; G. H. Martin, *Kinematics and Dynamics of Machines*, 2d ed., 1982; J. F. Shannon, *Marine Gearing*, 1981; J. E. Shigley and J. J. Uiker, *Theory of Machines and Mechanisms*, 2d ed., 1994.

Gecko

The name for about 300 species of reptiles that form the family Gekkonidae in the order Squamata. They are small lizards, primarily arboreal and nocturnal, which occur in the warm regions of the world. The body is flattened and never exceeds 14 in. (35 cm) in length. Most species have five digits, while some have only four; the toes often have adhesive pads to assist the animal when climbing on smooth surfaces.

Geckos feed on small animals, especially insects, and all species have a long sensitive tongue to aid in capturing their prey. In some species, the tail is used as a fat-storage organ. These animals move with agility and speed, and if caught by the tail, they can break it off by a spontaneous muscular contraction and rapidly regenerate a new one. Most species of gecko are oviparous; although they lay their eggs in protected areas, no particular care is given to them. *See* REGENERATION (BIOLOGY).

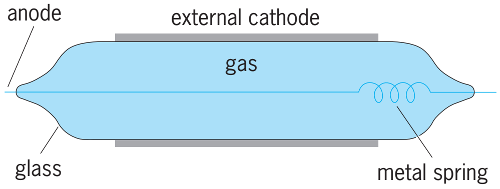
The largest, most aggressive species is the orange-spotted Tokay (*Gekko gecko*), which measures 14 in. (35 cm) and is indigenous to Southeast Asia. Another species is the flying gecko (*Ptychozoon homalocephalum*), which is about 8 in. (20 cm) long. It is essentially arboreal and is well adapted to leaping and gliding because it has folds of skin on either side of the body that can be opened out to form a planing surface. The banded gecko (*Coelonyx variegatus*) of the southwestern United States is about 3–4 in. (7.5–10 cm) long (see **illus.**). It is one of the few species with movable eyelids, and with claws instead of pads on the toes. *See* SQUAMATA. Charles B. Curtin

Banded gecko (*Coelonyx variegatus*).

Geiger-Müller counter

A detector of ionizing radiation. When a fast-moving charged particle traverses a Geiger-Müller counter, an electrical impulse is produced. These impulses can readily be counted by electronic circuits. Geiger-Müller (GM) counters, usually referred to simply as Geiger counters, are widely used to indicate the presence and intensity of nuclear radiations.

Construction and operation. A Geiger counter consists of a gas between two electrodes (see **illus.**). One electrode, usually cylindrical and hollow, is the cathode. The other electrode, a fine wire stretched along the axis of the cylinder, is the anode. A potential of about 1000 V is placed on the wire. When an atom of the gas between the two electrodes is ionized by collision with a charged particle passing through the gas, the electron produced in the collision is drawn toward the central wire. The electron



Cylindrical external-cathode Geiger-Müller counter, with thin soda glass and central wire of 0.003-in.-diameter (0.012-mm) tungsten. Metal spring keeps central wire taut.

then collides with the atoms of the gas. Near the central wire the electric field is very intense, and the electron may acquire enough energy between two collisions to allow it to ionize another atom. A second electron is then set free, and by successive collisions, an avalanche of electrons is produced which is then collected as charge on the central wire. This charge produces an electrical impulse which in typical cases may be 50 V.

A Geiger counter is a nonproportional particle detector; that is, the pulse produced is independent of the nature and energy of the particle that liberated the electrons within the gas. The operation of a Geiger counter depends critically on the voltage placed on the central wire. If the voltage is too low, the electron avalanche never builds up, and the counter operates only as an ionization chamber (a device which gives the total ionization produced in the gas) or a proportional counter, in which the output pulses are much smaller. If the voltage is too high, additional electrons produced by the positive ions start additional discharges, and the counter discharges continuously. Starting at the threshold for Geiger counting and continuing up to the region of multiple pulsing is the plateau region of the Geiger counter. The threshold is rather sharp, and occurs at an anode potential of 800–1000 V in typical counters. The plateau of a good Geiger counter extends 100–500 V above threshold. *See* IONIZATION CHAMBER; PARTICLE DETECTOR.

The pulses coming from a Geiger counter are amplified electronically and are then counted by an electromechanical register or as clicks in a loudspeaker. If the counting rate is so high that the mechanical register cannot follow the pulses fast enough, the pulses are fed into a scaling circuit, which divides the number of input pulses by a known scaling factor before they are fed to the register. The pulses may also be sent into a counting-rate meter, which contains a condenser that charges a certain amount each time a pulse reaches it. The accumulated voltage on the condenser indicates the counting rate. This voltage can be measured if a resistance is placed across the condenser to discharge it slowly.

Coincidence counting. Geiger counters are often used in coincidence. When a single particle passes through two or more Geiger counters, the pulses from each counter are practically time-coincident. The pulse from each counter is then sent to a coincidence circuit which indicates pulses coincident in time. Coincidence counting is used to eliminate much background count due to local radioactivity. Arrays of Geiger counters in coincidence also select particles going in a given direction. This technique is used, for example, to measure the angular distribution of cosmic rays.

Types and dimensions. Construction of Geiger counters requires careful control of both cleanliness of the electrode surfaces and purity of the gas. An inert gas, such as argon, is often used, mixed with a small amount of organic vapor to a total pressure of about 0.14 atm (2.1 lb/in.^{2} or 14 kilopascals). The organic vapor quenches the discharge and prevents multiple pulsing. However, some of the vapor is used during each pulse and such tubes thus have a limited life, usually 10^8–10^9 counts. Geiger tubes quenched by halogen gas are also available, and these have much longer life. The central wire of a Geiger counter is usually made of tungsten, 0.003–0.005 in. (75–125 micrometers) in diameter. The outer electrode may be copper, brass, or aluminum. For certain applications, external cathode counters, such as the one shown in the illustration, are used. The cathode is formed by spraying a conducting coating on the thin glass envelope surrounding the central wire. The electrical conductivity of soda glass, while small, is sufficient to permit the current to pass through the glass. Such counters have a long recovery time, but are useful in cosmic-ray research and are much cheaper than metallic counters. They have the additional advantage of not being destroyed if an overvoltage is accidentally applied.

Geiger counters may be made in various sizes, depending on the use for which they are intended. Counters 0.5 in. (13 mm) in diameter and 3 in. (75 mm) long, with thin aluminum walls, are used for detection of beta particles and cosmic rays. Counters of diameter 2 or 3 in. (50 or 75 mm) have been made, with lengths up to 3 ft (0.9 mm). However, if a large area is to be covered by GM counters, it is customary to use several connected electronically in parallel. Point Geiger counters have a somewhat different geometry than cylindrical counters. A sharp point is used as the anode, and a flat plane or sphere is the cathode.

Applications. Geiger counters are widely used in industry, medicine, mineral exploration, and scientific research.

Industry. A typical industrial use is in the measurement of thickness of sheet material. A source of beta rays is placed on one side of the material and a Geiger counter on the other. Some of the beta rays are absorbed in the material, and the counter responds to changes in absorption caused by changes in the thickness. Continuous accurate measurement of the thickness can be made even if the sheet material is moving at high speed. Another industrial application is in the detection of traces of radioactive elements in research on the wear of bearings. Such counters are used in connection with x-ray machines to search for flaws in large castings.

Medicine. Radioactive sources used in medicine are often detected with Geiger counters, which can locate the position or distribution of the radioactive material after it has been administered. In medical applications it is important to determine the total radiation dosage and to monitor the radiation; Geiger counters are used for this purpose. They have also been valuable in locating expensive and potentially dangerous sources of radioactivity.

Geology. Geologists use Geiger counters in mineral exploration. Minerals containing uranium are radioactive, and can be detected by Geiger counters. Simple lightweight portable Geiger detectors are available for prospectors. *See* GEOPHYSICAL EXPLORATION; RADIOACTIVE MINERALS.

Research. Chemists, physicists, and biologists use Geiger counters in various types of scientific research. Radioactive tracers used in chemical research are followed through complicated chemical reactions by Geiger counters. The counters give information on the quantity of material present and also on the lifetimes of radioactive products produced in the transmutation of elements. Geiger counters can be designed to employ a radioactive gas instead of the gas normally used. In such counters, the wall thickness is not important because the ionizing particle originates within the counter. Applications of Geiger counters in high-energy nuclear physics are confined to low-counting-rate experiments, such as in cosmic rays, because of the long dead time of Geiger counters. The counters are rarely used near particle accelerators, for they jam in a large flux of ionizing particles.

William B. Fretter

Bibliography. C. F. Delaney and E. C. Finch, *Radiation Detectors: Physical Principles and Applications*, 1992; K. Klein Knecht, *Detectors for Particle Radiation*, 1988; G. F. Knoll, *Radiation Detection and Measurement*, 3d ed., 1999; W. J. Price, *Nuclear Radiation Detection*, 2d ed., 1964.

Gel

A continuous solid network enveloped in a continuous liquid phase; the solid phase typically occupies less than 10 vol % of the gel. Gels can be classified in terms of the network structure. The network may consist of agglomerated particles (formed, for example, by destabilization of a colloidal suspension; **Fig. 1***a*); a "house of cards" consisting of plates (as in a clay) or fibers (Fig. 1*b*); polymers joined by small crystalline regions (Fig. 1*c* and **Fig. 2**); or polymers linked by covalent bonds (Fig. 1*d*).

In a gel the liquid phase does not consist of isolated pockets, but is continuous. Consequently, salts can diffuse into the gel almost as fast as they disperse in a dish of free liquid. Thus, the gel seems to resemble a saturated household sponge, but it is distinguished by its colloidal size scale: the dimensions of the open spaces and of the solid objects constituting the network are smaller (usually much smaller) than a micrometer. This means that the interface joining the solid and liquid phases has an area on the order of 1000 m^2 per gram of solid. As a result, the properties of a gel are controlled by interfacial and short-range forces, such as van der Waals, electrostatic, and hydrogen-bonding. Factors that influence these forces, such as introduction of salts or another solvent, application of an electric field, or changes in pH or temperature, affect the interaction between the solid and liquid phases. Variations in these parameters can induce huge changes in volume as the gel imbibes or expels liquid, and this phenomenon is exploited to make mechanical actuators or hosts for controlled release of drugs from gels. For example, a polyacrylamide gel (a polymer linked by covalent bonds) shrinks dramatically when it is transferred from a dish of water (a good solvent) to a dish of acetone (a poor solvent), because the polymer chains tend to favor contact with one another rather than with acetone, so the network collapses onto itself. Conversely, the reason that water cannot be gently squeezed out of such a gel (as from a sponge) is that the network has a strong affinity for the liquid, and virtually all of the molecules of the liquid are close enough to the solid-liquid interface to be influenced by those attractive forces. *See* HYDROGEN BOND; INTERMOLECULAR FORCES.

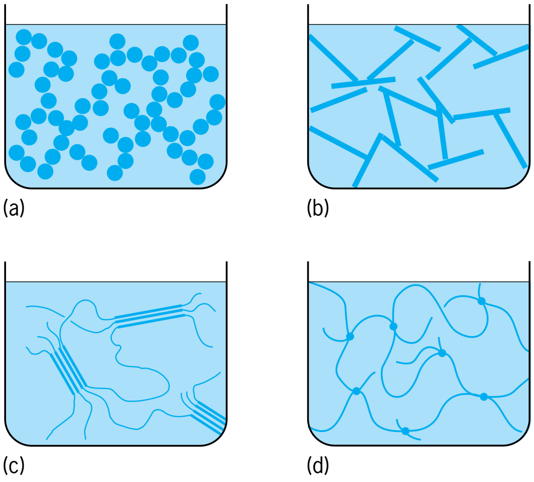


Fig. 1. Gel structures. (*a*) Agglomerated particles. (*b*) Framework of fibers or plates. (*c*) Polymers linked by crystalline junctions. (*d*) Polymers linked by covalent bonds. (*After M. Djabourov, Architecture of gelatin gels, Contemp. Phys., 29(3):273–297, 1988*)

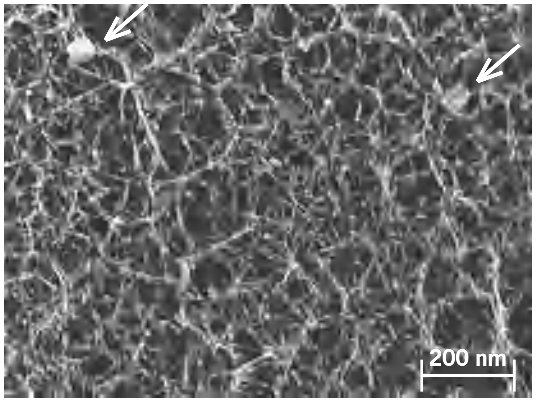


Fig. 2. Electron micrograph of a gelatin gel, where the solid phase occupies ~2% of the total volume. The arrows indicate latex particles with diameter of 90 nm that were included in the solution before gelation to serve as markers. The water was removed from the gel by vitrification (rapid quenching causing freezing without crystallization of the water) and sublimation. (*From M. Djabourov, Architecture of gelatin gels, Contemp. Phys., 29(3):273–297, 1988*)

Gelation. The most striking feature of a gel, which results from the presence of a continuous solid network, is elasticity: if the surface of a gel is displaced slightly, it springs back to its original position. If the displacement is too large, gels, except those with polymers linked by covalent bonds, may suffer some permanent plastic deformation, because the network is weak. The process of gelation, which transforms a liquid into an elastic gel, occurs with no change in color or transparency and no evolution of heat. It may begin with a change in pH that removes repulsive forces between the particles in a colloidal suspension, or a decrease in temperature that favors crystallization of a solution of polymers or the initiation of a chemical reaction that creates or links polymers. *See* PH.

When the particles in a colloid are attracted to one another, they aggregate into a cluster. If the particles form weak bonds when they collide, each particle will tend to detach from the cluster and reattach at another spot, seeking a place where it can bond to several other particles at once. This process leads to dense crystalloids or precipitates that settle out of solution. However, if each collision between particles results in an irreversible bond, the random nature of the collisions leads to a loose, disorganized aggregate with a great deal of occluded liquid. Collisions between clusters produce even less dense structures, until finally the whole volume of the original colloid is crowded with wispy aggregates jostling against one another, and gradually linking into a single giant cluster that becomes the network of the gel. In the early stage of this process, all of the bonds are formed within clusters that are free to diffuse about, so the colloid still flows, and is properly called a liquid. (This stage is accurately described by the theory of cluster-cluster aggregation.) As large clusters begin to link together, the viscosity of this liquid rises drastically. The gel point is reached when enough bonds have formed between clusters that a continuous network extends throughout the volume of the original colloid, arresting flow and imparting elasticity to the system. (The evolution of structure and properties in the vicinity of the gel point is properly described by percolation theory.) *See* AMORPHOUS SOLID; COLLOID; PRECIPITATION (CHEMISTRY).

Aging and drying. In many gelling systems the bonds are strong but not permanent, so that the structure of the gel changes slowly over time. This process is called aging. If the solid phase is soluble in the liquid, coarsening (termed Ostwald ripening) will occur, and sometimes an amorphous gel network will crystallize. In polymeric systems, flexible branches of the network may come into contact and form new bonds, leading to a gradual contraction of the network (known as syneresis).

In many cases, both the liquid and solid phases of a gel are of practical importance (as in timed release of drugs, or a gelatin dessert). More often (as in preparation of catalyst supports, chromatographic columns, or desiccants) it is only the solid network that is of use. If the liquid is allowed to evaporate from a gel, large capillary tension develops in the liquid, and the suction causes shrinkage of the network. The porous network remaining after evaporation of the liquid is called a xerogel, and it usually retains high porosity (>25%) and enormous surface area (150–900 m^2/g), along with very small pore size (1–10 nm). To maximize the porosity of the dried product, the gel can be heated to a temperature and pressure greater than the critical point of the liquid phase, where capillary pressures do not exist. The fluid can then be removed with little or no shrinkage occurring. The resulting solid is called an aerogel, and the process is called supercritical drying. Such materials are useful as Cerenkov radiation detectors, thermal insulation, and catalytic substrates. *See* DESICCANT; DRUG DELIVERY SYSTEMS; DRYING; GELATIN; HETEROGENEOUS CATALYSIS.

Sol-gel processing. Sol-gel processing comprises a variety of techniques for preparing inorganic materials by starting with a sol, then gelling, drying, and (usually) firing. Many inorganic gels can be made from solutions of salts or metallorganic compounds, and this offers several advantages in ceramics processing: the reactants are readily purified; the components can be intimately mixed in the solution or sol stage; the sols can be applied as coatings, drawn into fibers, emulsified or spray-dried to make particles, or molded and gelled into shapes; xerogels can be sintered into dense solids at relatively low temperatures, because of their small pore size. Moreover, hybrid materials can be made by combining organic and inorganic components in the gel. Many hybrids have such compliant networks that they collapse completely during drying, leaving a dense solid rather than a porous xerogel; therefore, they can be used as protective coatings (on plastic eyeglass lenses, for example) without heat treatment. Hybrid gels show great promise for active and integrated optics, because optically active organic molecules retain their activity while encapsulated in the gel matrix.

The most important application of gels continues to be for the preparation of thin films (on the order of a micrometer in thickness). Dense films, which must be sintered, typically at 500–700°C (932–1292°F), are used for antireflective or infrared reflective coatings, and protection against corrosion or abrasion. Unfired porous films, usually with encapsulated organic molecules, are being studied for use as sensors. Xerogel membranes are of interest for filtration and gas separation.

Catalytic substrates and chromatographic media have been made by sol-gel processing for decades, but improvements in control of microstructure should provide substantially improved performance. By growing the gel in the presence of surfactants, ordered mesoporous microstructures (consisting of regular cylindrical tubes, uniformly spaced layers, and even three-dimensional networks) have been prepared. These structures can be compared to zeolites, which have uniform pores on the order of 0.3–1 nm and which allow reactions to be performed exclusively on the molecules that fit into the pores; the mesoporous gels have pores an order of magnitude larger than zeolites, so size- and shape-selective chemistry can be extended to much larger molecules. Control over the microstructure of chromatographic media has been achieved by inducing phase separation in a solution of polymers and metalorganic compounds. After the inorganic gel network is formed, the polymer is removed by washing or thermal decomposition, leaving a uniform interconnected network of relatively large pores (up to the micrometer scale) within the micro- or mesoporous gel. The large pores provide rapid transport into the interior of the gel body, while the micropores perform the chromatographic separation.

The preparation of aerogels by supercritical drying was first reported in 1931, and there has been recurrent interest in their use as catalysts and as thermal insulation, but the cost of fabrication has so far prevented widespread use. It has been demonstrated that materials with the same physical structure as aerogels can be made without the expensive and hazardous supercritical processing. For example, if a silica gel is treated with chlorosilanes, it shrinks during conventional drying but then springs back to its original dimensions; alternatively, if a gel is aged in a solution containing additional silicate monomers, it becomes so rigid that it does not shrink during drying, so a low-density xerogel is obtained. *See* RHEOLOGY; SINTERING; SOL-GEL PROCESS. George W. Scherer

Bibliography. C. J. Brinker and G. W. Scherer, *Sol-Gel Science*, 1990; D. DeRossi et al. (eds.), *Polymer Gels: Fundamentals and Biomedical Applications*, Plenum, 1991; P. J. Flory, *Principles of Polymer Chemistry*, 1953; K. Nakanishi, Pore structure control of silica gels based on phase separation, *J. Porous Mat.*, 4:67–112, 1997; Proceedings of the 5th International Symposium on Aerogels, *J. Non-Crystalline Sol.*, vol. 225, 1998; Royal Society of Chemistry, *Faraday Discussions*, no. 101, *Gels*, 1995; Special issues of *J. Sol-Gel Sci. Technol.*, Organic/Inorganic Hybrids, 5(2), 1995, 7(3), 1996, and Non-linear Optical Materials, 9(2), 1997, and Optical Solids, 11(3), 1998; T. Tanaka, Gels, *Sci. Amer.*, 244:124–138, 1981; R. Zallen, *The Physics of Amrophous Solids*, 1983.

Gel permeation chromatography

A separation technique involving the transport of a liquid mobile phase through a column containing the separation medium, a porous material. Gel permeation chromatography (GPC), also called size exclusion chromatography and gel filtration, affords a rapid method for the separation of oligomeric and polymeric species. The separation is based on differences in molecular size in solution. It is of particular importance for research in biological systems and is the method of choice for determining molecular weight distribution of synthetic polymers.

Procedure. The separation medium is a porous solid, such as glass or silica, or a cross-linked gel which contains pores of appropriate dimensions to effect the separation desired. The liquid mobile phase is usually water or a buffer for biological separations, and an organic solvent that is appropriate for the sample and is compatible with the column packing for synthetic polymer characterization. Solvent flow may be driven by gravity, or by a high-pressure pump to achieve the desired flow rate through the column. The sample to be separated is introduced at the head of the column (**Fig. 1**). As it progresses through the column, small molecules can enter all pores larger than the molecule, while larger molecules can fit into a smaller number of pores, again only those larger than the molecule. Thus, the larger the molecule, the smaller is the amount of pore volume available into which it can enter. The sample emerges from the column in the inverse order of molecular size; that is, the largest molecules emerge first followed by progressively smaller molecules. In order to determine the amount of sample emerging, a concentration detector is located at the end of the column. Additionally, detectors may be used to continuously determine the molecular weight of species eluting from the column. The volume of solvent flow is also monitored to provide a means of characterizing the molecular size of the eluting species.

As the sample emerges from the column, a concentration detector signal increases above the baseline, passes through a maximum, and returns to the baseline (**Fig. 2**). This signal provides a relative concentration of emerging species and is recorded as a function of elution volume. Molecular-weight detectors based on light scattering or viscometry produce a signal, independent of molecular weight, that increases with increasing molecular weight for a given sample concentration. The molecular-weight chromatograms therefore are skewed with respect to the concentration detector signal. Typically, constant-volume pumps are used, and the time axis is transformed into a volume axis. Devices such as siphons that empty after a known volume of solvent has collected, or electronic circuitry that measures the transit time of a thermal pulse between two points in the

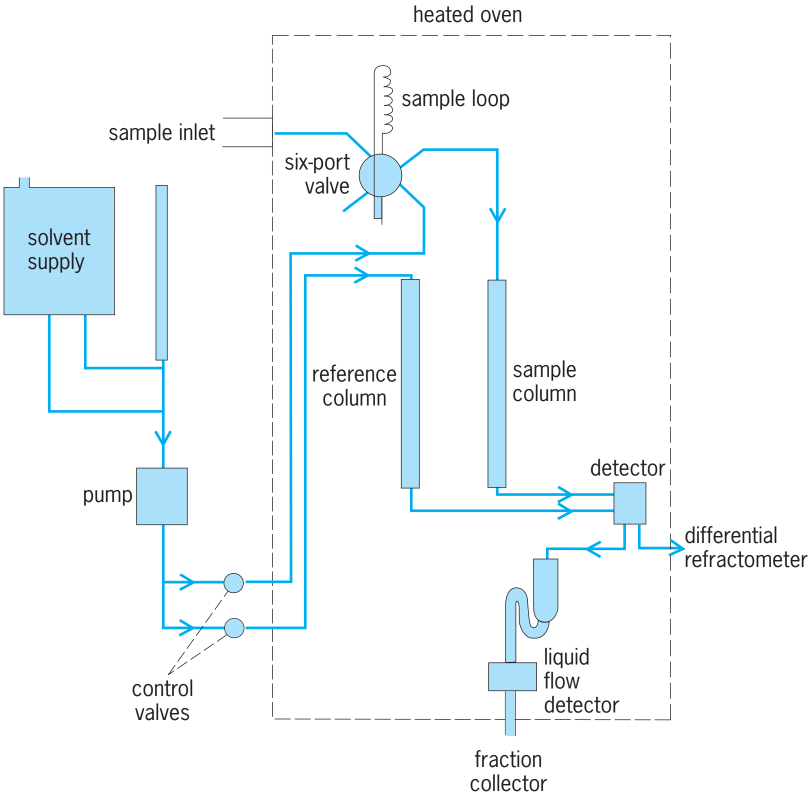


Fig. 1. Gel permeation chromatograph.

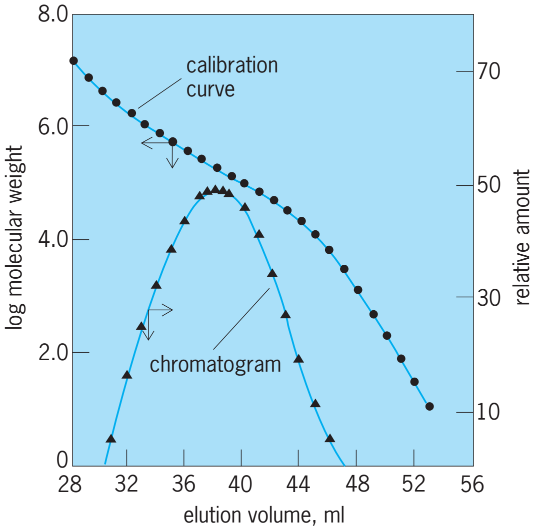


Fig. 2. Typical gel permeation chromatogram with a calibration curve.

flowing solvent, have been used to provide a measurement of the volumetric flow rate.

Calibration procedure. In the absence of direct molecular-weight detection, a calibration curve (Fig. 2) is required to convert elution volume into molecular weight. This is normally prepared by injecting standard samples of a polymer with a narrow molecular-weight distribution, so that the molecular weight at the maximum in the chromatogram may be unambiguously assigned. It must be assumed that elution volume depends only on molecular weight; that is, the only operative separation mechanism is size exclusion. The molecular weights of the standards are determined by primary molecular-weight methods, such as light scattering of dilute solutions or membrane osmometry, which determine the weight-average and number-average molecular weights, respectively. Secondary molecular-weight methods such as dilute-solution viscometry and vapor-pressure osmometry are also commonly used.

The procedure described above requires a set of characterized molecular-weight standards for each polymer type being characterized. Two approaches have been pursued to circumvent this problem. One approach involves an extensive search for a "universal" calibration curve, that is, for one on which a wide variety of molecular structures could be represented by a single line. It has been established that a plot of the logarithm of the intrinsic viscosity times the molecular weight versus elution volume produces such a curve. Many high-molecular-weight materials can be characterized from this curve, which is established by using only one type of material. This quantity, intrinsic viscosity times molecular weight, is proportional to the hydrodynamic volume. It appears, therefore, that the separation is based mainly on molecular size in solution. There are other contributing factors such as diffusion and nonsteric effects that affect separation, but these seem to be relatively unimportant compared to the size factor.

The other approach has been to develop detectors that continuously measure the molecular weight or some parameter that is proportional to molecular weight. The former involves the use of laser light scattering to measure the weight-average molecular weight. The latter is obtained from a continuously recording viscometer, measuring the intrinsic viscosity, which is proportional to the viscosity-average molecular weight. If the polymer eluting from the column at any time is essentially monodisperse, these two averages will be identical. The intrinsic viscosity may be converted into molecular weight by use of the Mark-Houwink relationship of intrinsic viscosity to molecular weight. Both approaches require the concurrent use of a concentration detector to measure the concentration of species eluting at a given time. A typical chromatogram obtained by using both light scattering and viscometric detectors in addition to a concentration detector has three curves (**Fig. 3**).

Light-scattering molecular-weight detectors allow calibration curves to be determined for a given column set by using either a set of narrow molecular-weight distribution fractions or a broad molecular-weight distribution polymer of undetermined molecular-weight distribution. Viscometric detectors allow a calibration curve and the universal calibration curve to be determined from either a set of narrow molecular-weight distribution fractions or a broad molecular-weight distribution polymer if the Mark-Houwink relationship is known. Similarly, the universal calibration curve may be determined by using a broad molecular-weight distribution polymer or narrow molecular-weight distribution fractions of a polymer of unknown chemical composition; or the Mark-Houwink relationship may be used if light scattering and viscometric detectors are used simultaneously. *See* VISCOSITY.

Once a calibration curve relating elution volume to molecular weight or the universal calibration curve is available, it is comparatively simple to convert the raw data from the chromatogram, that is, relative concentration versus elution volume, to a molecular-weight distribution. If molecular-weight detectors

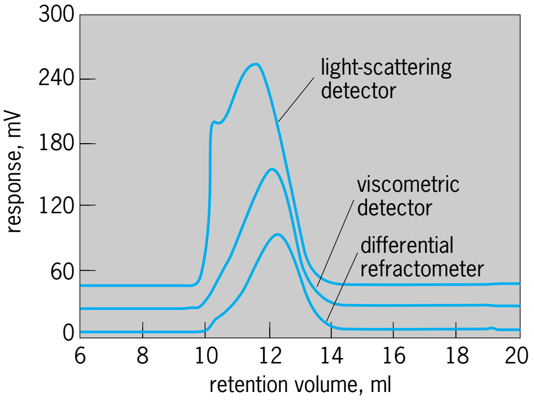


Fig. 3. Typical triple chromatogram obtained by using both light scattering and viscometric detectors in addition to a concentration detector.

are used, a direct relationship of molecular weight versus relative concentration is obtained. Such calculations are performed automatically by the data acquisition computer. The usual parameters calculated are the cumulative and differential molecular-weight distributions. The various molecular-weight averages may be calculated from these distributions.

Other parameters such as inhomogeneity factors and descriptions of the breadth of the distribution may be computed if desired. Thus, the formerly laborious process of determining precise molecular-weight distributions becomes very simple with the use of gel permeation chromatography. *See* MOLECULAR WEIGHT.

Column packings. A wide variety of porous materials have been used for column packing. For aqueous applications, cross-linked dextrans and polyacrylamide have found widespread use. In addition, cross-linked hydrophilic gels and inorganic materials with appropriate surface treatments having greater mechanical strength have been used. For organic solvents, cross-linked polymeric gels such as polystyrene and various other cross-linked copolymers are frequently employed. Rigid porous materials such as glass and silica that have high strength and extremely high solvent resistance have also been shown to produce effective separations. The trend has been to produce columns having column packings with particle diameters of a few micrometers. These columns tend to be fairly short to avoid large pressure drops across them, which would crush or distort the packing materials. Faster analyses are achieved, and the beneficial side effect is lower solvent use.

Detectors. In analytical applications, the concentration of solute is extremely small in order to maintain high-resolution separations and to avoid sample viscosity effects. Sample elution concentration is continuously monitored by using differential refractometry, infrared and ultraviolet/visible spectroscopy, colorimetric detectors, and various combinations of these. Multiple detectors are required if the samples are not chemically homogeneous. Some of these detectors are extremely sensitive to temperature and flow rate, so variations must be kept to a minimum. *See* COLORIMETRY; REFRACTOMETRIC ANALYSIS; SPECTROSCOPY.

Associated equipment. Pumping systems range from columns using gravity flow to high-pressure precision pumping systems that can be regulated to give a constant flow rate with very good accuracy. Soft gel column packings will not tolerate high pressures, and slower analyses result. Samples may be injected by hypodermic syringe or from calibrated volume injection loops attached to multiport valves. Columns may be metal or glass and have diameters from a few millimeters to several centimeters for preparative columns, and lengths from a few centimeters to several meters. In general, column efficiency increases directly with column length, while band spreading increases as the square root of column length. Thus, in principle, with a long enough column it should be possible to separate very difficult resolvable materials. The time for analysis, however, increases linearly with column length.

Flow rate. As the analysis time is inversely proportional to flow rate, it is advantageous to use the highest permissible flow rate that will give the required resolution. The effect of flow rate has been extensively studied. Column performance may be defined in terms of the number of plates per unit length as measured with monodisperse (essentially homogeneous) substances. In general, it is found that efficiency decreases with increasing flow rate. The decrease, however, is not large, and in typical analytical columns of approximately 0.2 in. (5 mm) inside diameter where flow rates are of the order of 1–5 milliliters per minute, a decrease of only a few percent is noted. At higher flow rates, the decrease in efficiency becomes more pronounced. Flow-rate limitations are sometimes imposed by the pressure at which a given column set may be operated. High flow rates should be avoided for high-molecular-weight samples because of viscosity effects and the possibility of shear degradation of the sample.

Sample size. Normally, the sample size for analytical gel permeation chromatography is 1 mg or less. As the molecular weight increases, the sample size should be reduced.

Solvent and temperature. Selection of the solvent and temperature is normally based on two factors. The first is that the sample must be soluble and nonassociated in solution. The second is compatibility with the detector system. For the differential refractometer and light-scattering detectors, a large difference between the refractive index of the sample and solvent is desirable. For spectrophotometric detectors, the absence of solvent absorption in the desired wavelength range is required. For viscometric detectors, use of a thermodynamically good solvent will cause a higher signal at a given molecular weight. A thermodynamically good solvent is one in which the polymer coil is expanded and is reflected in a higher value of a certain constant in the Mark-Houwink relationship.

Unlike most chromatographic separations, temperature has very little effect on separability, as the separation is based on the hydrodynamic volume of the molecule. There are minor effects on this volume with changes in temperature. Normally, however, all the species will change in the same direction and approximately the same amount so there is little overall effect on the relative separation. There are minor changes in efficiency due principally to changes in viscosity. Lower viscosity of the solvent improves mass transfer and gives a faster approach to equilibrium. The changes due to viscosity are usually also minor; therefore, the temperature chosen is normally just above ambient so that adequate temperature control of the solvent can be achieved to permit steady baselines from the detectors used. For some synthetic polymers, for example, the polyolefins, of difficult solubility, it is necessary to operate at elevated temperatures, and usually antioxidants must be added to the solvent to prevent oxidative degradation of sample and solvent. Temperatures ranges from below ambient to 150°C (300°F) have been reported.

Applications. Gel permeation chromatography has had widespread applications. For determining the molecular weight of synthetic polymers, at least 50 types of polymers have been characterized. These include alkyd resins, natural and synthetic rubbers, cellulose esters, polyolefins, polyamides, polyesters, polystyrenes, polyacrylates, uncured epoxy, urethane and phenolic resins, and a wide variety of oligomeric materials. Additionally, the ability to determine the molecular-weight distribution and changes in distribution has led to many applications in areas such as blending distributions, chain-length studies in semicrystalline polymers, interactions in solution, radiation studies, mechanical degradation studies, mechanisms of polymerization research, polymerization reactor control, and evaluation of the processing of polymers. In the field of natural and biological polymers, numerous systems have been separated and analyzed. Among these are acid phosphatases, adrenalin, albumin, amino acids and their derivatives, enzymes, blood group antibodies, collagen and related compounds, peptides, and proteins.

Additionally, gel permeation chromatography is capable of making separations of low-molecular-weight compounds. This is particularly important when both low- and high-molecular-weight species are present in the same sample. *See* CHROMATOGRAPHY; POLYMER. Anthony R. Cooper

Bibliography. S. Mori and H. G. Barth, *Size Exclusion Chromatography*, 1999; B. H. Stuart, *Polymer Analysis*, 2002; C.-S. Wu (ed.), *Handbook of Size Exclusion Chromatography and Related Techniques*, 2d ed., 2003.

Gelatin

A protein extracted after partial hydrolysis of collagenous raw material from the skin, white connective tissue, and bone of animals. It is a linear polymer of amino acids, most often with repeating glycine-proline-proline and glycine-proline-hydroxyproline sequences in the polypeptide linkages. Gelatin gives positive test for hydroxyproline, and yields amino acid components upon complete hydrolysis. *See* AMINO ACIDS; COLLAGEN; PROTEIN.

Characteristics. The unique characteristics of gelatin are reversible sol-to-gel formation, amphoteric properties, swelling in cold water, film-forming properties, viscosity-modifying properties, and protective colloid properties. The last is exemplified by a very low Zsigmondy gold number. Gelatin solutions are levorotatory; the specific optical rotation, which greatly increases at lower temperatures, is −137° at 589 nanometers and 45°C (113°F). *See* OPTICAL ACTIVITY.

Most physical properties of gelatin are measured on aqueous solution, in either the sol or gel state. The physical rigidity or gel strength of gelatin is measured at 6.67% concentration with a Bloom gelometer and is expressed in grams. Gelatin from first extracts yields 250–325 Bloom gelatin, whereas the last extract might yield 50–100 Bloom gelatin. Viscosity of 6.67% gelatin solution at 140°F (60°C) can range from 25 to 65 millipoises, with most commercial gelatin being in the 40–50 millipoise range. The pH is usually adjusted to between 5.2 and 6.5, depending on the end-use specifications. Gelatin contains molecular species from 15,000 to 400,000, with the average molecular weight between 50,000 and 70,000 for 200–300 Bloom gelatin.

Chemical composition. Gelatin contains 26.4–30.5% glycine, 14.8–18% proline, 13.3–14.5% hydroxyproline, 11.1–11.7% glutamic acid, 8.6–11.3% alanine, and in decreasing order arginine, aspartic acid, lysine, serine, leucine, valine, phenylalanine, threonine, isoleucine, hydroxylysine, histidine, methionine, and tyrosine.

Absence of only two essential amino acids—tryptophan and methionine—makes gelatin a good dietary food supplement. Most of the active groups in gelatin molecules are either terminal or part of side groups. An α chain with a molecular weight of about 96,500 behaves in a solution like a random-coil polymer, whereas the gel form may contain as much as 70% helical conformation. With the same gelatin, the rigidity of the gel is approximately proportional to the square of the gelatin concentration. There are two types of α chain, with slight chemical differences, designated α_1 and α_2. The original collagen triple helix contains two α_1 and one α_2 chain. The dimer made up from either $\alpha_1 + \alpha_2$ or $\alpha_1 + \alpha_1$ in gelatin is called β, and a trimer with a molecular weight of 298,500 is called γ. Homogeneous α-chain gelatin has been prepared by pretreating collagen with pronane in the presence of 0.4 *M* calcium chloride ($CaCl_2$). Enzyme-pretreated gelatin has the isoionic point of type A gelatin.

Manufacturing process. Collagen, the precursor of gelatin, is pretreated for 10–30 h with 1–5% mineral acid for type A gelatin production, or 35–90 days with a lime slurry for type B gelatin production. Type A gelatin exhibits the isoionic point at pH 7.0–9.5, whereas type B gelatin, due to deamination in the liming process, exhibits the isoionic point at 4.8–5.2. Under alkaline conditions, arginine is slowly hydrolyzed to ornithine by liberating urea. A short, 1–3-day alkaline process using 10–15% solution of sodium hydroxide (NaOH) and sodium sulfate (Na_2SO_4) has been patented in Great Britain. Type A gelatin is manufactured from frozen and fresh edible-grade pig skins or from bone ossein after demineralization. A type A gelatin made from hides has also become available commercially. Bovine hides and skins (known as splits) and trimmings of raw or salted dehaired hide pieces are used for type B gelatin. Some type B gelatin is also made from pig skins in Japan and in China. Most of type B gelatin, however, comes from bones. The bone pieces (0.5–2 cm) are first demineralized in 4–7% hydrochloric acid for 7–14 days before liming. After pretreatment with either acid (type A) or lime (type B), the materials are washed and then are subjected to four or five 4–8-h extractions at increased temperatures of 131–212°F (55–100°C). The extracts, containing 3–7% gelatin, are filtered, concentrated in vacuum evaporators, chilled and extruded as noodles, and dried at 86–140°F (30–60°C). The dry gelatin is then ground and blended to specifications. The viscosity of consecutive extracts from bone ossein increases, whereas, the trend in type A gelatin from pig skins is reversed.

Uses. The principal uses of gelatin are in foods, pharmaceuticals, and photographic industries. Other uses of industrial gelatin are in the field of microencapsulation, health and cosmetics, and plastics.

Foods. The food industry is the largest user of gelatin. The formulations use exclusively either water or aqueous polyhydric alcohol solutions. The principal uses are in gelatin desserts, marshmallows, candies, meat products such as canned hams and luncheon meats, canned soups, bakery icings and pie fillings, sour cream, and ice cream. Flavors and vitamins can be encapsulated in gelatin to control release or to protect against aerial oxidation. Gelatin is also recommended to counteract fingernail defects. Most edible gelatin is of type A, but type B gelatin is also used.

Photography. Gelatin has been used in photography for over 100 years as a binder in light-sensitive and backing layers. After 1960 the photographic industry switched to so-called inert type B gelatin, where chemical sensitizers and restrainers have been removed and the gelatin no longer controls the silver halide crystal growth or its final light sensitivity. The function of gelatin remains as a protective colloid, as a bromine acceptor, and as a matrix which allows coating and processing of modern photographic formulations. Manufacture and quality control of inert photographic gelatin is very critical, since impurities of the order of parts per million can have a pronounced photographic effect. The largest-volume photographic products are x-ray films, phototypesetting paper, color paper, lithographic films, and microfilm. Derivatized gelatins such as phthalated gelatin also find wide use in the photographic market. Type A gelatin use in photography is limited to graphic arts products. *See* PHOTOGRAPHIC MATERIALS.

Pharmaceuticals. The largest use of gelatin in pharmaceuticals in the United States is in the manufacturing of hard and soft capsules for medications. The formulations are made with water or with aqueous polyhydric alcohols. The gelatin for soft capsules is a low-Bloom type A (6.0–6.3 oz or 170–180 g), type B (5.3–6.2 oz or 150–175 g), or a mixture of the two. The material is encapsulated between two plasticized gelatin sheets, and shaped and sealed in a rotary die. Hard capsules consist of two tightly fitting halves prefabricated from medium- to high-Bloom type A (8.8–9.9 oz or 250–280 g), type B (7.9–8.8 oz or 225–250 g), or the mixture of the two types. After the capsule-tampering scare in the 1980s, a different form of gelatin-coated medication referred to as gelcaps was introduced in addition to caplets. Tamper-proof capsules have been introduced also. A special pyrogen-free gelatin grade is used as a blood plasma substitute. Glycerinated gelatin is used in suppositories. Enteric capsules are made with a cross-linked or coated gelatin. Absorbable gelatin sponge (Gelfoam) is gelatin foam partially insolubilized by cross-linking, and is used for arresting hemorrhage during surgery.

Technical applications. In the plastics industry gelatin can be used to control polymerization rate and particle size. Electroplating in gelatin solution produces a bright finish to electroplated surfaces of silver, gold, copper, brass, and zinc. Gelatin coating can protect copper, zinc, and aluminum surfaces against corrosion, and is used as a pickling inhibitor for stainless steel. Gelatin can be a source for nitrogen, as well as time-release coating for fertilizer components. In animal feed, gelatin is used to stabilize and pelletize vitamins and minerals otherwise susceptible to aerial oxidation or degradation. A different application of gelatin is its use for simulating body tissue consistency in the study of ballistics. A great deal of research has been done in this field with ordnance-type gelatin. *See* ANIMAL FEEDS; ELECTROPLATING OF METALS; FERTILIZER; POLYMERIZATION. Felix Viro

Bibliography. Gelatin Manufacturers Institute of America, Inc., *Gelatin*, 1983; L. Stryer, *Biochemistry*, 4th ed., 1995.

Gem

A mineral or other material that has sufficient beauty for use as personal adornment and has the durability to make this feasible. With the exception of a few materials of organic origin, such as pearl, amber, coral, and jet, and inorganic substances of variable composition, such as natural glass, gems are lovely varieties of minerals. *See* JET (GEMOLOGY).

Natural Gems

A mineral is defined as an inorganic substance with a characteristic chemical composition and usually a characteristic crystal structure. Each distinct mineral is called a species by the gemologist. Two stones that have the same essential composition and crystal structure but that differ in color are considered varieties of the same species. Thus ruby and sapphire are distinct varieties of the mineral species corundum, and emerald and aquamarine are varieties of beryl. Two or more minerals that have the same structure and are related chemically are called groups. For example, almandite, pyrope, rhodolite, grossularite, andradite, and some other minerals form the garnet group. *See* MINERAL; MINERALOGY.

Most gemstones are crystalline (that is, they have a regular atomic structure) and have characteristic properties, many of which have a direct bearing on the gem's beauty or durability.

Durability. Each mineral has a characteristic hardness (resistance to being scratched) and toughness (resistance to cleavage and fracture).

Cleavage. Cleavage is the tendency of a crystalline substance to separate parallel to a crystal face or a possible crystal face. The term is also applied to such a separation visible in a crystalline substance.

Fracture. Fracture is a break or a tendency to break in a direction that bears no relationship to the atomic structure of the mineral.

Hardness. The hardness of a gem material determines its resistance to being scratched and thus, in conjunction with resistance to cleavage and fracture, its practicability for use as a jewelry stone.

The hardness of minerals is measured with reference to an empirical scale known as Mohs scale, which consists of 10 minerals, ascending from talc (hardness 1) to diamond (hardness 10). The scale was formulated by determining which of the various well-known minerals scratched others. Gypsum (2) scratches talc, calcite (3) scratches gypsum, and so on up the scale. However, the hardness of diamond is much greater in relation to corundum (9) than is corundum in relation to topaz (8). With few exceptions, the most important gemstones are those at the top of the hardness ladder; for example, diamond is 10, ruby and sapphire are 9, topaz and beryl (emerald and aquamarine) are 8, and the quartz family has a hardness of 7. *See* HARDNESS SCALES.

Beauty. Optical properties are particularly important to the beauty of the various gem materials. The important optical properties include color, dispersion (or "fire"), refractive index, and pleochroism. Gemstones usually are cherished for their color, brilliancy, fire, or one of the several optical phenomena, such as the play of color of a fine opal or the star effect in a sapphire.

Brilliancy. Brilliancy depends on the refractive index, transparency, polish, and proportions of a cut stone. Refractive index is a measure of a gem's ability to reduce the velocity of light that passes through it. It is defined either as the ratio of the velocity of light in air to the velocity of light within a gemstone, or as the ratio of the sine of the angle of incidence to the sine of the angle of refraction. Although diamond does not have the highest refractive index known, it has one of the highest to be found among transparent materials. This quality, combined with its unsurpassed hardness (which permits it to be given a magnificent polish) and great transparency, gives diamond the highest brilliancy potential of any mineral or manufactured material. Although synthetic rutile (titania, or titanium dioxide) has a higher refractive index, its lower transparency, relative softness (which makes a superior polish impossible), and other factors, such as its large birefringence, combine to reduce its brilliancy well below that of diamond. *See* REFRACTOMETRIC ANALYSIS.

In order for a gemstone to display its latent brilliancy to best advantage, it must be cut to the proper angles for its refractive index. For example, the optimum angle of the pavilion (bottom portion) of a brilliant-cut diamond is 41°, whereas the stones of lower refractive index should be cut to slightly greater angles (measured from the plane of the girdle).

Refractive indices. The refractive indices of the important gems vary from a low of approximately 1.45 for opal to a high among transparent stones of 2.6–2.9 for synthetic rutile. The value for diamond is 2.42. Refractive index is a rough measure of brilliancy; in other words, the brilliancy of two colorless stones of equal transparency and cutting quality is approximately in proportion to their refractive indices. The refractive indices of other important gems are as follows: ruby and sapphire, 1.76; spinel, 1.72; chrysoberyl, 1.74; topaz and tourmaline, 1.62; beryl, 1.58.

Dispersion. Dispersion is the breaking up of white light into its component colors. It is measured by the difference in the refractive indices of certain of the red and blue wavelengths of light as they pass through a gem. In order to give comparable figures for the dispersive powers of different gemstones, two specified wavelengths in the red and blue ends of the Sun's spectrum are used. For example, the refractive indices for these two wavelengths in spinel are 1.710 and 1.730, respectively; therefore, the amount of dispersion is expressed as 0.020. A comparable figure for synthetic rutile is 0.330.

Selective absorption. Although diamond and colorless zircon are valued for their brilliancy and prismatic fire, gemstones are more often prized for the loveliness of their colors. The color of most gemstones is caused by selective absorption. Selective absorption refers to a gem's ability to absorb or transmit certain wavelengths of light more readily than others. The wavelengths transmitted with least absorption are those that give a gem its color. Absorption is usually caused by very small percentages of metallic oxides that are present as impurities.

Pleochroism. Some gemstones exhibit pleochroism, the property of some doubly refractive materials of absorbing light unequally in the different directions of transmission, resulting in color differences. If two different colors are exhibited, the result is called dichroism (displayed by ruby, sapphire, and emerald); if three colors are seen, the result is called trichroism (displayed by alexandrite and tanzanite). *See* DICHROISM; TRICHROISM.

Asterism. Some gemstones depend on unusual optical effects for their beauty. Perhaps the most important of these are the star stones, which display the phenomenon called asterism. In a star, the reflection of light from lustrous inclusions is reduced to sharp lines of light by a domed cabochon style of cutting. The usual star effect is the six-rayed star seen in star sapphires and rubies. Four-rayed stars occur in garnets and some spinels. Other six-rayed stars are occasionally seen in such stones as quartz and beryl. In corundum, the most important species in which asterism is a significant feature, the phenomenon is caused by reflection from three sets of needlelike inclusions arranged in planes perpendicular to the sides of the hexagonal (six-sided) crystal, with each of the three sets of inclusions in planes parallel to a pair of faces of the crystal. *See* CORUNDUM.

Chatoyancy. Another important optical phenomenon is chatoyancy, or a cat's-eye effect. The most important gem in which it is a prominent feature is the variety of the mineral chrysoberyl known as precious cat's-eye. This effect, when seen in chrysoberyl, is usually much more silkily lustrous than in any other gemstone. As a result, the term precious cat's-eye has come to be applied to the finer specimens of the chrysoberyl variety. Other kinds of gemstones in which cat's-eyes are sometimes encountered include tourmaline, quartz, beryl, scapolite, diopside, and some of the other rarer gem minerals.

Play of color. The gem mineral opal is one of the most attractive of the phenomenal gems. The colorful display called play of color results from light interference caused by diffraction effects from tiny spheres of material joined to make up the body of the opal.

Adularescence. Another optical phenomenon that produces an interesting gem variety is adularescence, the billowy light effect seen in adularia or moonstone varieties of orthoclase feldspar. In its most attractive form, it has a slightly bluish cast. Moonstone and opal are among the softer of the fairly common gem minerals and have hardnesses of approximately 6 on Mohs scale.

Precious and semiprecious stones. For many years gemstones have been commonly designated as precious or semiprecious. This is a somewhat meaningless practice, however, and often misleading, since many of the so-called precious gem varieties are inexpensive and many of the more attractive varieties of the semiprecious stones are exceedingly expensive and valuable. For example, a cabochon of fine-quality jadeite may be valued at approximately 100 times the price per carat of a low-quality star ruby. Fine black opals, chrysoberyl cat's-eyes, and alexandrites are often much more expensive than many sapphires of certain colors. *See* CARAT.

Occurrence. Gemstones occur under a variety of conditions in nature. A number of the highly prized minerals occur as primary constituents in igneous rocks or in alluvial deposits. The only known occurrences of diamonds in primary deposits are in either of the ultrabasic rocks (those with a low percentage of silica—less than 40%) known as kimberlite, a type of peridotite; or in lamprophyre. Although most primary diamond deposits are volcanic necks or conduits (so-called pipe deposits), dikes and sills of kimberlite or lamprophyre also may contain diamonds. Diamond-bearing pipes are known in South Africa, Tanzania, the two Congos, Angola, Sierra Leone, India, Lesotho, Russia, Australia, Canada, and the United States (Arkansas). Extensive alluvial deposits are mined in Ghana, Angola, Zaire, Sierra Leone, Guinea, Ivory Coast, Namibia, South Africa, Brazil, Guyana, and Venezuela. *See* IGNEOUS ROCKS; LAMPROPHYRE; PERIDOTITE.

Perhaps the most important types of primary deposits in which gemstones other than diamond are found are pegmatite dikes, contact metamorphic deposits, and weathered lava flows containing corundum. The intrusion of molten rock into impure limestone often brings about the crystallization of calcite to form a coarse-grained marble, plus a broad assemblage of other minerals. Under ideal conditions, some of the minerals produced in the contact zone include gem-quality crystals; for example, corundum and spinel. The Mogok area of Myanmar is the foremost supplier of the finest rubies and very fine sapphires. Pegmatite dikes are found in a number of places throughout the world, but only a few of them produce gem-quality crystals. The most important gem pegmatite areas include Brazil, Madagascar, Mozambique, Zambia, Namibia, Nigeria, and southern California. Minerals recovered from pegmatite dikes include aquamarine (and the other varieties of beryl other than fine emerald), topaz, tourmaline, and the kunzite variety of spodumene. Fine emeralds occur in calcite veins in black shale near Muzo, Colombia. Elsewhere most emeralds occur in mica schist, sometimes with other beryllium minerals, such as chrysoberyl. *See* PEGMATITE.

Most of the important colored-stone mining is done in alluvial deposits. For example, the island of Sri Lanka, long known as the "island of gems," has produced a variety of gemstones from alluvial deposits for many years.

Identification. Since most gems are varieties of mineral species, and since mineral species have fairly constant chemical compositions and characteristic crystal structures, their physical and optical properties vary within rather narrow limits. Thus, identification is a matter of measuring these properties and other characteristics, the most important of which are refractive index, optic character, and specific gravity. Because less than 10% of all known mineral species have varieties that are useful as gems, comparatively few tests are needed to separate them; only the first three tests are required in the majority of identifications. However, the ready availability of synthetic gemstones, plus a number of different enhancement techniques to improve the appearance of natural gemstones, complicates this task.

Origin. The properties of synthetic gem materials are identical to those of their natural counterparts; therefore, it is usually necessary to study suspected stones under magnification to determine whether their growth characteristics are natural or the result of synthetic production. For example, the accumulation lines in the form of color banding in natural sapphires are arranged in a hexagonal pattern, whereas those in many synthetic materials are curved. Spherical gas bubbles are characteristic of synthetics, but the inclusions in natural sapphires are angular. These same conditions also hold true for ruby and synthetic ruby. To make this determination, binocular magnification with dark-field illumination is the most effective (**Fig. 1**).

Refractive index. The relative refractivity of a gemstone is measured on a refractometer, an instrument that measures the critical angle between an optically dense hemisphere and the gemstone being tested. The instrument is calibrated directly into refractive indices and, as a result, readings from the scale are not in angles but in refractive indices. Doubly refractive minerals show two readings if the birefringence is sufficiently pronounced (birefringence is the difference between the two indices of a doubly refractive material). Thus, a characteristic birefringence

Fig. 1. Gemolite, a binocular magnifier with dark-field illuminator. (*Gemological Institute of America*)

Fig. 2. Polariscope to distinguish between single and double refraction. (*Gemological Institute of America*)

may be read as well as a characteristic refractive index. For example, tourmaline shows two readings (one at 1.624, another at 1.644), which alone are sufficient to identify this gem mineral. If monochromatic sodium light is used instead of ordinary incandescent illumination, it is possible to read indices accurately to the third decimal place. When two readings are seen, the necessity for the use of the polariscope, the instrument used to distinguish between single and double refraction, is unnecessary. The polariscope (**Fig. 2**) utilizes two polaroid plates, set so that their vibration directions are at 90° to one another, and the gemstone is examined between the two plates. A singly refractive gemstone rotated between the plates remains uniformly dark, whereas a doubly refractive stone becomes alternately light and dark. Diamonds, garnets, and glass remain dark in the polariscope, whereas emerald, ruby, tourmaline, topaz, and zircon become alternately light and dark.

Specific gravity. The density of a gem material compared with an equal volume of water is measured either by heavy liquids or by weighing the stone in air and in water on a sensitive balance adapted for this purpose. Gem materials vary in specific gravity from amber, at about 1.05, to hematite, at 5.20.

Fluorescence characteristics in ultraviolet light, spectroscopy, x-ray diffraction, chemical composition, and hardness and other tests may be helpful in identification (see **table**).

Gem materials. More than 100 natural materials have been fashioned at one time or another for ornamental purposes. Of these, however, only a relatively small number are likely to be encountered in jewelry articles.

Hardness, specific gravity, and refractive indices of gem materials

Gem material	Hardness	Specific gravity	Refractive index
Amber	$2-2\frac{1}{2}$	1.05	1.54
Beryl	$7\frac{1}{2}$.–8	2.67–2.85	1.57–1.58
Synthetic emerald	$7\frac{1}{2}-8$	2.66–2.7	1.56–1.563 to 1.57–1.58
Chrysoberyl and synthetic	$8\frac{1}{2}$	3.73	1.746–1.755
Corundum and synthetic	9	4.0	1.76–1.77
Diamond			
Synthetic cubic	10	3.52	2.42
Zirconia	$8\frac{1}{2}$	5.80	2.15
Feldspar	$6-6\frac{1}{2}$	2.55–2.75	1.5–1.57
Garnet			
Almandite	$7\frac{1}{2}$	4.05	1.79
Pyrope	$7-7\frac{1}{2}$	3.78	1.745
Rhodolite	$7-7\frac{1}{2}$	3.84	1.76
Andradite	$6\frac{1}{2}-7$	3.84	1.875
Grossularite	7	3.61	1.74
Spessartite	$7-7\frac{1}{2}$	4.15	1.80
Hematite	$5\frac{1}{2}-6\frac{1}{2}$	5.20	
Jade			
Jadeite	$6\frac{1}{2}-7$	3.34	1.66–1.68
Nephrite	$6-6\frac{1}{2}$	2.95	1.61–1.63
Lapis lazuli	5–6	2.4–3.05	1.50
Malachite	$3\frac{1}{2}-4$	3.34–3.95	1.66–1.91
Opal	$5-6\frac{1}{2}$	2.15	1.45
Pearl	3–4	2.7	
Peridot	$6\frac{1}{2}-7$	3.34	1.654–1.690
Quartz			
Crystalline and synthetic	7	2.66	1.54–1.55
Chalcedony	$6\frac{1}{2}-7$	2.60	1.535–1.539
Spinel and flux synthetic	8	3.60	1.718
Synthetic spinel, flame	8	3.64	1.73
Spodumene	6–7	3.18	1.66–1.676
Topaz	8	3.53	1.61–1.63
Tourmaline	$7-7\frac{1}{2}$	3.06	1.62–1.64
Turquois	5–6	2.76	1.61–1.65
Zircon	$7\frac{1}{2}$	4.70	1.925–1.98
Metamict	7	4.00	1.81
Zoisite (tanzanite)	6–7	3.35	1.691–1.70

Amber. Amber is a fossil resin that is usually yellow to brown, but it is occasionally found in other colors. In gem quality it is used mostly for beads or carvings. Amber is also used for pipe stems, cigarette holders, and other items. *See* AMBER.

Beryl. Beryl is a beryllium-aluminum silicate. In gem quality it is usually transparent and faceted, but translucent beryl may be carved into attractive forms used in jewelry. When green, it is known as emerald; when light blue, as aquamarine; when pink to light red, as morganite. It also occurs in yellow, light green, brown, orange, medium red, and colorless forms. *See* BERYL; EMERALD.

Chrysoberyl. Chrysoberyl is a transparent to translucent beryllium aluminate. Its varieties, known as cat's-eye and alexandrite, are among the important gemstones. Cat's-eye exhibits a band of light across the dome of a cabochon, which is caused by reflections from parallel silklike inclusions. The color is greenish yellow to yellowish green. Alexandrite changes color from green in daylight to garnet red in candlelight. *See* CHRYSOBERYL.

Coral. The gem material coral is an assemblage of colonies of the tiny marine animal coral. It is usually an orange to flesh color, but may be white or deep brownish red, and semitranslucent to nearly opaque.

Corundum. Corundum is transparent to translucent aluminum oxide. Ruby is the orange-red to violet-red variety; all other colors are called sapphire. The principal colors of sapphire other than blue are yellow, colorless, orange, pink, purple (amethystine), and green. Both ruby and sapphire occur with a beautiful six-rayed star effect. *See* RUBY; SAPPHIRE.

Diamond. Diamond, the transparent form of carbon, is the most important gemstone. In addition to the popular colorless form, it occurs in yellow, brown, pink, blue, and green. Bombardment by subatomic particles (followed by heat treatment for some colors) yields green, yellow, orange, and brown; blue is also possible. *See* DIAMOND.

Feldspar. The gemstone of importance in the feldspar group is moonstone. It is a semitransparent form of potassium feldspar that exhibits a floating light effect. Less important gems in the feldspar family include the amazonite variety of microcline, sunstone (a spangled gold-colored variety resembling moonstone, except for color), and the iridescent, translucent gray labradorite. Since several minerals are included in the feldspar group, the properties vary from species to species. *See* FELDSPAR; LABRADORITE; MICROCLINE; ORTHOCLASE.

Garnet. Within the garnet group are several distinct minerals that share the same crystal structure but differ in physical properties and chemical composition. Two of the garnet species, almandite and pyrope, have the dark red color commonly ascribed to garnet. A mixture of the almandite and pyrope compositions in about a 3:1 ratio produces a distinctively colored stone that is given the separate name rhodolite. It is a lighter, more transparent gemstone, with a distinctive violet-red color. Among other species of garnet are andradite, of which the lovely green demantoid is a variety; grossularite, of which the orange-brown hessonite and the green tsavorite are varieties; and spessartite, which has attractive orange and orange-red varieties. *See* GARNET.

Hematite. The principal ore of iron, hematite, sometimes occurs in dense, hard masses that take a very high polish; these are often fashioned into intaglios and cameos. Hematite is metallic grayish black in color. *See* CAMEO; HEMATITE; INTAGLIO (GEMOLOGY).

Jade. The material popularly known as jade may be a variety of either jadeite, a member of the pyroxene group, or nephrite, a combination of amphibole group minerals. Pyroxenes and amphiboles are major rock-forming minerals. The jades are characterized by a felted structure, creating exceptional toughness, which makes even delicate carvings relatively durable. In their gem qualities, they are also prized for their attractive coloring and translucency. Because it may occur in more saturated colors and may be more nearly transparent, green jadeite potentially is the more expensive of the jade minerals. It is semitransparent to semitranslucent and may be intense green, white with green streaks or patches, mauve, brown, yellow, orange, or violet. Nephrite is translucent to opaque and usually is a darker, less intense green than jadeite; it may also be off-white, gray, or black. *See* JADE.

Lapis lazuli. Lapis lazuli is an opaque, vivid deep-blue gemstone flecked with golden-yellow pyrite. It has been used for many centuries and is undoubtedly the stone to which the term "sapphire" was first applied. *See* LAZURITE.

Malachite. The colorful, opaque mineral malachite is often banded in two or more tones of green and may show a radial fibrous structure. It is always green and is often accompanied by the deep violet-blue azurite, another copper mineral. Malachite is used mostly for cameos and intaglios and for inexpensive scarab bracelets. *See* AZURITE; MALACHITE.

Opal. The fabled beauty of opal is due to diffraction and light interference in which patches of intense colors are seen against either a white or a nearly black background. There is also a transparent orange to red variety called fire opal. Since it is a relatively fragile material, opal must be treated with care. *See* OPAL.

Pearl. Oriental pearls are those found as lustrous, nacreous concretions in one of the three species of the salt-water mollusk genus *Pinctada*. Concretions in edible oysters are without pearly luster and are valueless. Pearls found in several genera of fresh-water clams are called fresh-water pearls, to distinguish them from Oriental pearls. Pearls usually occur in white, cream, or yellow colors with rose or other overtones, although black and gray are also very desirable. *See* PEARL.

Peridot. Peridot is a yellowish-green to green variety of the mineral group olivine; it is a magnesium-iron-aluminum silicate. In gem quality, it is always transparent and olive green. *See* OLIVINE.

Quartz. In its two major types, single crystal and cryptocrystalline (or chalcedonic), quartz, the most common mineral, has more gem varieties than any other mineral species. The important varieties of crystalline quartz are amethyst, citrine, aventurine, and tiger's-eye; the most important varieties of cryptocrystalline quartz are carnelian, sard, chrysoprase, bloodstone, agate, and onyx. Amethyst is purple to violet and transparent. Citrine is yellow to brown and also transparent; it is more commonly known as topaz-quartz and, unfortunately, is often sold as a topaz. Aventurine is usually green with lustrous or colored spangles; it is translucent. Tiger's-eye is a translucent, fibrous, broadly chatoyant, yellow-brown stone that may be dyed other colors. Carnelian is red to orange-red, and sard is a darker brownish red to red-brown; both are translucent. Chrysoprase is light yellowish green. Bloodstone is dark green with red spots. Agate is a translucent chalcedony showing curved bands in a variety of colors. Onyx is similar to agate, except that the bands are straight. *See* AGATE; AMETHYST; ONYX; QUARTZ.

Spinel. This gem mineral is of particular interest because of the strong resemblance of many of its varieties to comparable colors of corundum. In general, red spinel is less intense in color than ruby; similarly, blue spinel is less intensely colored than blue sapphire. On the other hand, some varieties of spinel are lovely in their own right. Flame spinel is an intense orange-red and is a very attractive gemstone; spinel also occurs in green and amethystine colors. All gem varieties of spinel are transparent, with the exception of the rare black star spinel. *See* SPINEL.

Spodumene. Spodumene is a member of the pyroxene group of minerals; in contrast to jadeite, however, it is fragile. The principal variety of spodumene for gem purposes is kunzite, which is a lovely light-red to light-purple transparent stone. *See* SPODUMENE.

Topaz. Topaz is best known in its yellow to brown variety, but the red, pink, and blue varieties are also attractive and desirable. Colorless topaz can be irradiated in a nuclear reactor and then heat-treated to produce an attractive light blue color. *See* TOPAZ.

Tourmaline. Although the color range of tourmaline is as wide as that of any gem material known in nature, its best known varieties are red (rubellite) and dark green; it may also be colorless, yellow, blue, black, brown, or other colors. *See* RUBELLITE; TOURMALINE.

Turquoise. Turquoise is an opaque gemstone with an intense light-blue color. Its intense color has attracted people from the earliest times. *See* TURQUOISE.

Zircon. Zircon is best known as a transparent colorless or blue gemstone. The colorless variety has been used principally as an inexpensive substitute for diamond. It also occurs in green, yellow, brown, red, and flame colors. The properties of zircon vary rather widely. *See* ZIRCON.

Zoisite. A mineral that had been associated with rather drab varieties achieved a gem quality when, in 1967, in Tanzania, a variety dubbed tanzanite was discovered. It has an exceptionally pleasing vivid blue color, reminescent of the finest so-called Kashmir sapphires from the Himalaya area.

Manufactured Gems

Some minerals or other materials that have sufficient beauty and durability for use as a personal adornment can be manufactured. The term "manufactured," as used here, does not include such processes as shaping, faceting, and polishing, but only the processes that affect the material from which the finished gem is produced. These processes are (1) those that change the mineral in some fundamental characteristic, such as color, called a treated gem; (2) those by which a material is made that is identical with the naturally occurring mineral, called a synthetic gem; and (3) those that produce a simulated material with the appearance but not both the composition and structure of the natural gem, called an imitation gem. *See* PRECIOUS STONES.

Treated gems. Gem treatment to improve color or to conceal or make flaws less obvious has become increasingly pervasive. Although the U.S. Federal Trade Commission mandates full disclosure, treatments are not revealed. Some minor treatments are undetectable, such as gentle heating to change or make a color more uniform. Major treatments have a profound effect on apparent value. If the separations on a heavily flawed diamond are filled with a foreign substance with a high refractive index, the fractures may become all but invisible. Later heating during a prong retipping, or ring sizing may make them more visible than before the treatment, or cause additional damage.

In a different context, many colored stones, including most green tourmaline, aquamarine, and colorless and flame-colored zircon and most pink topaz and blue zircon have been heated to improve their color. This tends to be regarded in the gem industry as being as much a part of normal market preparation as cutting and polishing. This does not apply to intense heating of close to the melting point such as is done to impart color by diffusion of a coloring agent into sapphire, for example.

There are a number of different types of treatment, including dyeing or staining, impregnation by paraffin, plastic, or some other material; filling in fractures or cavities to hide them; bleaching; laser drilling to reach ugly inclusions in order to make them less obvious; heat treatment; diffusion of color into the near surface of a gem too light in color to be high in value; and irradiation by subatomic particles. In some efforts to make a stone more attractive, two or more treatments are employed.

Dyeing and staining. This method of coloration was one of the earliest to have been adopted widely. The most frequently used gem subjected to dye is agate (cryptocrystalline quartz, or chalcedony). Colored substances are introduced into the minute interstices between the microscopic quartz crystals. Because of the varying porosity in the individual layers, the difference in color between adjacent layers is intensified. The best method of coloring a given specimen is usually determined by trial-and-error treatment of a small sample. In general, the process of dyeing an agate consists of first drying the specimen thoroughly; second, soaking it in an appropriate liquid dye until the interstices are penetrated; and finally, fixing the dye within the pores. In the processing of cultured pearls eosine dyes are often used.

Fracture and cavity filling. By a process first developed in Israel, diamonds with prominent cleavages or fractures are subjected to a treatment by which a liquid with a high refractive index is introduced into the separation. When it solidifies, the high refractivity of the transparent in-filling reduces the visibility of the separation. Thus, the clarity appears often to be improved greatly. Cavities in colored stones of high per-carat value are occasionally filled with a glass of similar refractive index as the host material. Many more emeralds contain inclusions than most other gemstones. Almost as long as emeralds have been used as gems, miners have presented them to buyers after soaking in some type of oil, which tends to improve apparent clarity.

Bleaching. Oxidizing agents or acids are sometimes used to remove unwanted colors. In the processing of cultured pearls employing bead nuclei, the first deposition is likely to be brown. After harvesting, this unwanted color gradually fades, but the processor speeds the fading by using a bleaching step to make matching for necklaces possible more quickly, and with less chance of later change in appearance.

Laser drilling. A diamond with a large and obvious internal inclusion that does not reach the surface may be enhanced in apparent clarity grade by employing a laser to drill a tiny hole to the feature, so that a bleaching or dissolving agent may reach it to reduce its visibility. Dramatic improvement in appearance is possible. *See* LASER.

Impregnation. The gemstone most often subjected to impregnation has been turquoise, but jade has also been bleached to remove an unattractive color, and then impregnated with plastic to restore durability. Since the demand for fine-quality natural turquoise greatly exceeds the supply, porous, chalky, poor-quality turquoise is soaked in paraffin or impregnated with plastic. The soaking deepens and improves color, and a plastic impregnation improves both color and durability.

Heat treatment. Heating in one or more different atmospheres is often used to improve the natural color of a gem or to develop new colors. As found in nature, most aquamarine is light to medium yellowish green, a color much less in demand that the lovely blue of a fine aquamarine. Light heating effects a permanent color change to the blue associated with aquamarine. Some amethyst becomes more evenly colored with gentle heating. When heated to about 450°C (842°F), most amethyst changes to a yellow- to reddish-brown citrine color, but the amethyst from a few localities changes to a green color when heated to this temperature, and the brown color of smoky quartz may often be changed to yellow or reddish-yellow citrine. Some yellow or brownish zircons change to blue, golden, or red when they are heated in a reducing atmosphere, or become colorless or rich yellow if heated in an oxidizing atmosphere. Over the years, heating techniques and controls have become increasingly sophisticated, with temperature controlled to less than 1°C (1.8°F), and with the type of atmosphere also controlled.

Diffusion. After effecting atmospheric and temperature controls to improve heat-treatment consistency, a logical next step was to introduce coloring agents. This practice has become very common, particularly in Bangkok. Faceted colorless or pale sapphires are heated to close to the melting point in a bed of the appropriate metallic oxide chromophore or chromophores. When successful, enough of the chromophore migrates into the lattice to impart color. This process may be repeated several times to increase the penetration.

Radiation. Irradiation of different types induces many color changes in gems. The induced color varies with the original material and with the kind and intensity of radiation used. X-rays and radiation from radium, radioactive isotopes, electrons, fast neutrons, and other sources may produce striking changes. Diamond subjected to bombardment in a cyclotron or a nuclear reactor first changes to a green color, and, with more radiation, the green passes to black with an accompanying breakdown in the crystal structure. If the irradiation is not carried too far, the change can be reversed to a rich yellow or brown by annealing at high temperatures. After annealing, the color changes appear to be permanent. Many changes induced by irradiation are short-lived. The term metamict is used to describe crystals in which structural changes have been introduced by intensive natural radiation. *See* METAMICT STATE; PLEOCHROIC HALOS.

Synthetic gems. The U.S. Federal Trade Commission has restricted the term synthetic gems to manufactured materials that have the same chemical, physical, and optical properties as their naturally occurring counterparts. Of the gem materials, including diamond, that have been synthesized, many have been produced in such small sizes or poor quality that they are unsatisfactory as gemstones. Some attempts to make gemstones have resulted in producing substances not known in nature, a few of which are of great importance industrially. Others have resulted in significant improvements in existing processes.

In 1891 E. G. Acheson, while attempting to synthesize diamond, produced SiC (carborundum), the third hardest known substance after diamond, and artificially produced borazon, a boron nitride. Knowledge gained in attempting to make ruby and sapphire, gem varieties of the mineral corundum (Al_2O_3), aided in the development of nongem crystals of corundum, which, with SiC, form the basis of the abrasives industry. The quantity of synthetic ruby used as jewel bearings in watches and electrical instruments greatly exceeds that used as gems.

In 1957 began the production and sale of an industrial-diamond grit for abrasive use, averaging

0.004 in. (0.1 mm) in diameter, in competition with natural diamond grit, which is made by crushing low-grade diamond crystals. The abrasive use of diamond grit for the manufacture of bonded-diamond-grinding wheels greatly exceeds in volume, but not in value, the amount of diamond used for other industrial applications and for gemstones. In 1971, the experimental production of cuttable-quality diamonds in sizes of over 1 carat was announced. These were grown under conditions of such extreme temperatures and pressures that the equipment costs made the product more costly than comparable natural diamonds. Since then, larger cuttable-quality crystals have been grown. In 1993, Russian-grown synthetic diamonds in cut stones of up to a carat were made available.

Verneuil process. This process uses the high temperatures produced by an oxyhydrogen flame. In 1902 A. V. L. Verneuil, a French chemist, announced the production of a synthetic ruby of gem quality made by a process that still bears his name; it is also called the flame-fusion method.

1. *Sapphire and ruby*. Synthetic ruby and sapphire are produced by the fusion of very finely divided and highly pure (Al_2O_3) in an apparatus that resembles an oxyhydrogen torch, with the flame directed downward into an insulating chamber. The powder passing through the flame fuses into droplets, which initiate the development of a cylindrical boule when the flame impinges on a fire-clay support. The diameter of the boule is controlled by lowering the fire-clay support while regulating the gas flow and the amount of Al_2O_3 introduced. The upper surface of the boule remains molten. Various colors are produced by the addition of appropriate metallic oxides. It is essential to exclude traces of sodium, which prevent the formation of clear boules. The average commercial Verneuil boule is about $^1/_2$ in. (13 mm) in diameter and 1–2 in. (25–50 mm) long, and weighs about 125 carats. Crystallographically oriented boules may be produced with a properly oriented seed crystal. Rods of transparent Al_2O_3 as small as 0.04 in. (1 mm) in diameter and up to 18 in. (46 cm) long may be made by a semiautomatic process. After being ground to uniform diameter, several rods are held firmly together in a bundle and sawed perpendicular to their length with a diamond saw. Each saw cut produces a circular blank from each rod, which is processed into a jewel bearing. From mushroomlike boules as large as 4 in. (10 cm) in diameter, circular disks are cut and polished. Because of the resistance of aluminum oxide to high temperatures, these disks are used as windows in furnaces. They are also highly conductive in the infrared region of the spectrum; and because of this conductivity, the rods may be inserted into furnaces to transmit the infrared radiation into heat-controlling apparatus on the outside of the furnace.

Ruby is used for jewel bearings in preference to clear sapphire because colored jewels are easier for workers to see. Ruby is made by adding a small percentage of chromium oxide (Cr_2O_3) to the alumina. Where extreme hardness is not essential, the more easily worked spinel is preferred. Colored varieties are easier to make and give clearer crystals than corundum for all colors except red. The hardness of spinel, 8 on Mohs scale, is adequate for most gem substitute purposes. Synthetic spinel boules are elongated cubes with rounded corners. Synthetic star rubies are made by adding an excess of titania (titanium oxide; TiO_2) to the alumina powder and then heating the finished boules at a temperature above 1830°F (1000°C). The titania recrystallizes as needles oriented by the host crystal, in the plane of the basal pinacoid, in a hexagonal pattern that is controlled by the crystal structure of the alumina.

2. *Rutile*. Another mineral synthesized by the flame-fusion process is rutile. For the production of synthetic rutile (titania), the flame-fusion apparatus is modified to oxidize the titanium more completely by the addition of a third tube at the orifice surrounding the other two tubes through which oxygen passes. The boules, as produced, are black and must be heated in a furnace in an oxygen atmosphere at 2550–2730°F (1400–1500°C) to remove the color. Unless metallic coloring oxides have been added, the boules retain a faint tinge of yellow after the oxygen-furnace treatment, because of an absorption band in the blue. The unoxidized titanium atoms in undercolorized boules make them semiconductors whose conductivity is proportional to the amount of color remaining. Synthetic titania boules form elongated tetragonal prisms with rounded corners. The boules are clear and transparent, larger than natural rutile crystals, and much superior in quality. Natural rutile is always brownish, even when transparent. Since the introduction of synthetic cubic zirconia, synthetic rutile is used rarely as a diamond substitute. *See* RUTILE.

3. *Strontium titanate*. Another once popular diamond substitute is strontium titanate. Strontium titanate made by the flame-fusion method has, since 1955, been a successful imitation of diamond. It has an index of refraction nearly the same as that of diamond and a somewhat higher dispersion, which gives it a brilliancy approaching, and a degree of fire somewhat greater than, diamond.

4. *Other melt methods*. These include the Czochralski process, the floating zone process, and the skull melt. Czochralski, or crystal pulling technique, melts the appropriate material, such as aluminum oxide in an inert crucible, inserts a seed crystal barely into the melt and withdraws it gradually as the melted material crystallizes on the retracting seed. By this method, large fine crystals of synthetic ruby, sapphire, alexandrite and other gem-quality crystals are produced. With modifications of this technique, very large crystals have been produced.

In the floating zone process of synthesis, a rod is prepared by sintering the essential ingredients of a gem material. The rod is held at both ends, and one end is subjected to a narrowly focussed convergence of infrared rays that bring the mixture to the melting point. The ends of the rod are rotated slowly in opposite directions. The heat is applied to move slowly down the rod, and it slowly becomes a single crystal as crystallization takes place from top to bottom. This is used to create synthetic rubies, sapphires, and the alexandrite variety of chrysoberyl.

The ingenious skull melt method for growing crystals of highly refractory materials owes its name to the appearance of the container. A series of water cooled tubes make up a container that is filled with the purified powder of the refractory material. The tubes are encircled by coils attached to a radio-frequency generator that heats the center portion of the powder to a very high temperature. The powder next to the cooling tubes insulates the tubes. This is the method by which huge quantities of zirconium oxide, the cubic zirconia (diamond substitute) is grown.

Flux growth. This is another fundamental method of synthetic crystal manufacture. Flux fusion, or flux melt, employs a material with a relatively low melting point in which the material to be synthesized will dissolve much more readily than in an aqueous solution. Slow cooling brings about the crystallization of the desired product.

Hydrothermal process. Synthesis by crystallization from aqueous solutions has been employed in numerous laboratories experimentally. It was first used on a large scale by the Bell Laboratories to produce large, colorless quartz crystals. Later, they made synthetic ruby experimentally. The first major gem synthesis by this method was announced by the Linde Company in 1965 in the making of synthetic emeralds. Hydrothermal synthetic emeralds are manufactured commercially.

Assembled stones. Once one of the most important imitations, the doublet of garnet and glass is of minor importance today. A combination of two parts of synthetic spinel, plus a colored cement, is still used extensively to imitate ruby, emerald, and several other gemstones. As assemblage of natural and synthetic ruby or sapphire is used to simulate star ruby or star sapphire. The only other important assembled stone is made of a thin piece of opal cemented to a backing of quartz or common opal and covered with a protective layer of transparent quartz.

Imitation gems. Since prehistoric times glass has been the most widely used gem imitation. Since World War II colored plastics have replaced glass to a great extent in the least expensive costume jewelry.

Glass. Brilliants and rhinestones made of flint glass are often silvered on the back to give them a high brilliancy. Glass with a high dispersion, which is often called paste diamond, has a high lead-arsenic content. The term paste owes its origin to the practice of wet mixing the ingredients before firing. It is somewhat softer than ordinary glass.

Plastics. Modern plastics have made costume jewelry very popular, not only because plastics are cheap but because they are lighter in weight than glass and are easily molded and worked into various forms. Thermosetting plastics, such as Bakelite, were first used in jewelry early in the twentieth century. Unlike thermoplastics, which can be resoftened by heating after molding, thermosetting plastics become permanently hard after the initial molding. *See* PLASTICS PROCESSING.

Identification of substitutes. Materials made by a flame-fusion process almost always contain gas bubbles, which are usually spherical or nearly so. Those with medium to dark tones of color often show color banding, or striae, with a curvature corresponding to that of the top of the boule. Natural gem materials are characterized by angular inclusions and straight color bands, if any are present. Often, there are differences in absorption spectra, fluorescence, or transparency to ultraviolet or other properties sufficient to enable a gemologist to identify them.

Most flux-fusion synthetic emeralds have distinctly lower refractive indices and specific gravities than natural emeralds and are characterized by wisplike or veillike flux inclusions. Most show a red fluorescence under ultraviolet light, whereas most natural emeralds are inert. Two- or three-phase inclusions (liquid and gas or solid, liquid, and gas enclosed in the same space) are typical of natural emerald.

Synthetic flux rubies have veillike inclusions and a transparency to shortwave ultraviolet slightly greater than that of natural rubies.

Hydrothermally made synthetic emeralds have properties similar to many natural emeralds, but their inclusions differ and they are characterized by a very strong red fluorescence under ultraviolet.

The cheaper forms of glass are cast in molds and, under a hand lens, show rounded edges at the intersections of facets; the facets are often concave. The better grades, known as cut glass, have been cut and polished after first being molded approximately into the desired form. Cut glass has facets that intersect in sharp edges. Both types may contain gas bubbles or have a roiled appearance in the heart or have both, in contrast to most of the colored stones they imitate.

Assembled stones are detected usually by observing the plane where the parts are joined together. In a faceted stone, these parts may be seen more readily if the stone is immersed in a transparent liquid and viewed parallel to the girdle.

Gem Cutting

The polishing of rough gem materials into faceted or rounded forms for use in jewelry is called lapidary or gem cutting. The term lapidary is limited in application to the cutting of colored stones. One who fashions colored stones is called a lapidary, lapidarist, or lapidist.

Commercial colored stone and diamond cutting are separate and distinct fields. The problems of colored-stone cutting are different from diamond cutting, because all colored stones are soft enough to be shaped readily by using silicon carbide as an abrasive, although diamond may be used for faster results. In contrast to diamond, they may be sawed or polished in any direction. There are major differences between the cutting of the most valuable emerald and ruby rough material, and the less expensive materials. Weight loss in fashioning is usually greater in colored stone cutting, because diamond shapes are usually more adaptable to the round

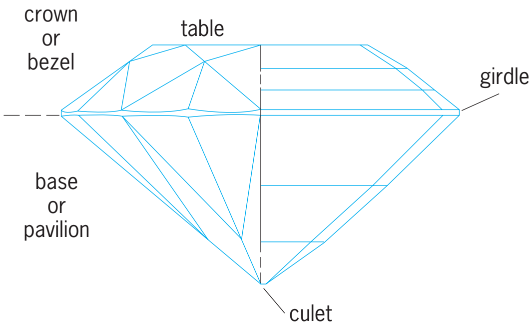


Fig. 3. Arrangement of faceted cuts, showing brilliant cut (left side) and step cut (right side).

brilliant that is the usual cut. Given the pleochroism of most valuable colored stones, hexagonal prisms are less likely to be adaptable to a high weight retention. Thus, when cutting the most valuable ruby and emerald crystals, planning is vital. Face-up appearance has to be balanced with weight retention potential. With less expensive rough material, mechanical, computer-controlled cutting may make it possible to cut dozens of exactly sized, almost cloned, stones at one time. In medium to expensive rough, an experienced lapidary may set by eye the angles to which facets should be cut in relation to the girdle plane or the top facet. In addition, the cutter creates by eye the perfect symmetry seen when the finished stone is viewed from above. In this respect, commercial gem cutting differs markedly from that done by hobbyist gem cutters, for the hobbyist usually controls the angles and the symmetry by the use of various mechanical faceting heads.

The first step in the usual cutting operation is to saw the crystal or piece of massive gem material to obtain the size and general shape desired. If the material is relatively inexpensive, it may be ground to shape rather than reduced by cutting. Next, if the stone is soft enough, it is usually ground into the approximate final shape against a silicon carbide wheel; if it is 9 on Mohs scale, a diamond-charged metal lap, or wheel, is used.

Facets are ground and polished on other laps. The commercial cutter often employs what is called the jamb-peg method. In this method, a square or rectangular board containing rows of partly drilled holes is clamped into a position at right angles to the spinning lap. The stone that is being faceted is mounted with adhesive on the end of a stick, called a dopstick, the opposite end of which is pointed. The cutter holds the pointed end of the stick in a hole chosen by the cutter as one representing the proper angle for the facet being ground. The end of the stick on which the stone is mounted is held against the spinning lap. The term dop is used both in diamond and colored-stone cutting.

Facet grinding is usually accomplished on copper, iron, or lead laps, using diamond or silicon carbide powder for ruby and sapphire or their synthetic counterparts, and silicon carbide powder for the softer materials. Polishing is accomplished on tin, wood, or plastic laps, with the wooden laps usually covered with leather or cloth. The polishing agents used include tin oxide, cerium oxide, alumina, and diamond powder. Both polishing and grinding laps are operated at rates of speed that are more rapid for hard materials than for soft ones.

Not all lapidaries appreciate the importance of the direct relationship between the proportions of the cut stone and its beauty. Rough gem material is so expensive that the natural tendency to retain every milligram possible in the final product often overcomes consideration of potential beauty. Since diamonds of fine quality are valued for their brilliancy and prismatic fire, the need for cutting to precise angles and proportions is appreciated by most diamond cutters. However, many depart somewhat from ideal figures to obtain larger stones that they believe will be more readily sold at a lower per-carat price. Since

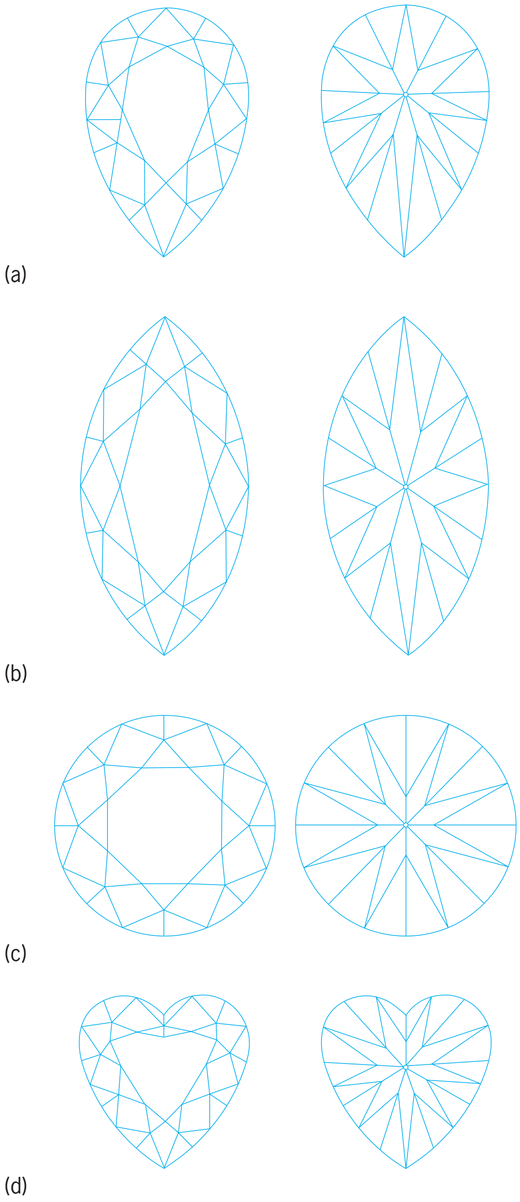


Fig. 4. Examples of brilliant-cut facets. (*a*) Pear-shape cut. (*b*) Marquise or navette cut. (*c*) Round brilliant. (*d*) Heart-shape cut. (*Gemological Institute of America*)

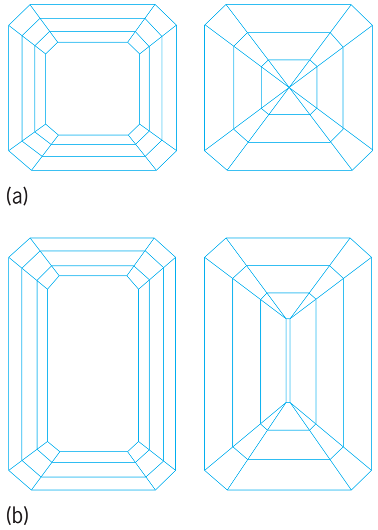


Fig. 5. Examples of step-cut facets. (*a*) Square emerald cut. (*b*) Emerald cut. (*Gemological Institute of America*)

colored stones are valued for their beauty of color, some lapidaries believe that any symmetrical cutting is satisfactory. Actually, the lower refractive indices of the important colored stones reduce both their potential brilliancy and their selective absorption, which produces the color. Many colored stones are so poorly cut that their reflective quality is seriously impaired.

There are many different styles to in which gems are cut. The two basic types are brilliant and step cutting. Both are characterized by a large facet (table) topping the crown facets, which slope to the periphery of the stone (girdle), and by a base portion, which slopes from the periphery to a tiny facet (culet) at the lowest point (**Fig. 3**). The two styles differ in that the facets on step cuts are parallel to the table and the girdle, so that they are all trapeze-shaped; whereas brilliant-cut facets, other than the table and culet, are triangular or kite-shaped. The brilliant style (**Fig. 4**) is usually used for stones with round or curved girdle outlines, and the step cut (**Fig. 5**) for those with square, rectangular, or other straight-sided girdle outlines.

Gem Mounting

Gemstones are usually set in rings or other jewelry pieces made of one of the precious metals. Those metals most widely used for jewelry purposes are yellow gold, white gold, platinum, and palladium.

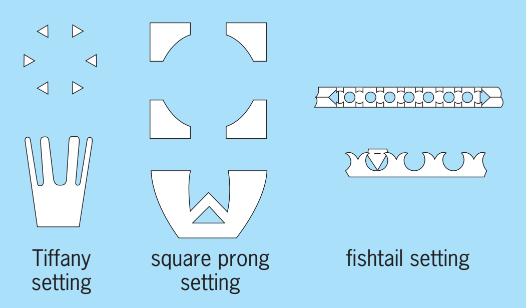


Fig. 6. Gemstone settings. (*Gemological Institute of America*)

Setting. When a stone is set in a ring, the portion in which the stone is placed is called the setting (**Fig. 6**). The stone may be held by small beads of metal, which are pushed over the edges by the stonesetter; by prongs, which are also bent over the edges of the stone to hold it securely; or by a rim of metal, which the stonesetter turns over the edges all the way around the stone.

Seat. In any type of setting in which the stone is held in place by metal, it is essential to the preservation of a stone subjected to normal wear that the "seat" into which it is placed be even. The seat is usually cut into the metal of the setting, after which the stone is pushed into position and the metal beads, prongs, or metal bezel are forced down around the edge of the stone. If the seat is uneven, the pressure on the stone is applied unevenly, and even a diamond may cleave or fracture when subjected to the light blows that it is certain to receive in the course of normal wear. Richard T. Liddicoat, Jr.

Bibliography. Gemological Institute of America, *Gem Reference Guide*, 1993; C. S. Hurlbut and R. C. Kammerling, *Gemology*, 2d ed., 1991; R. T. Liddicoat, *Handbook of Gem Identification*, 12th ed., 1987, reprint 1993; R. Webster and B. W. Anderson, *Gems: Their Sources, Description, and Identification*, 5th ed., 1994.

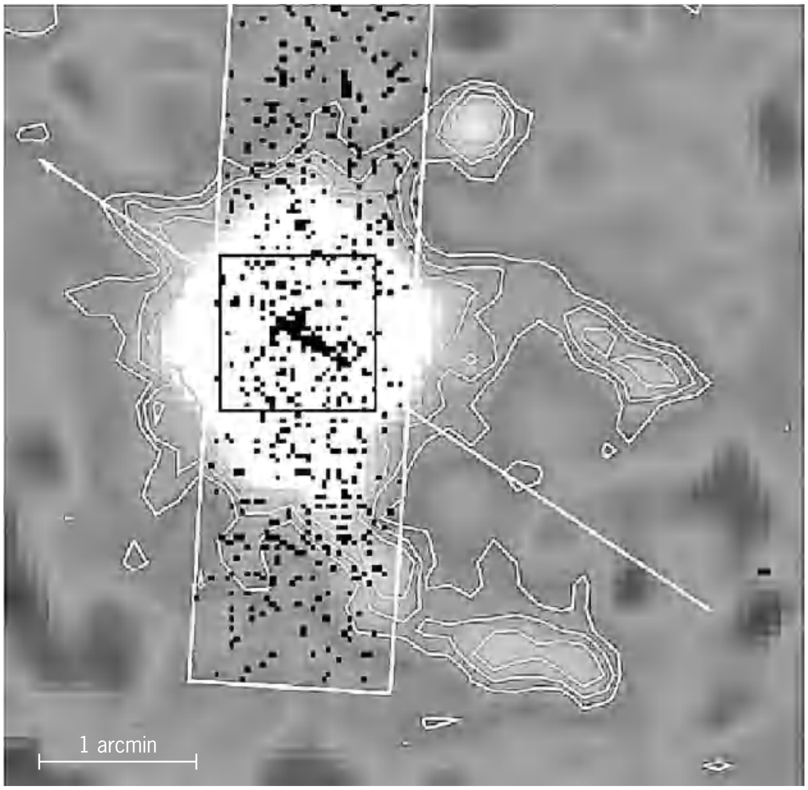
Geminga

A nearby neutron star that emits pulsed x-rays and gamma rays, steady optical radiation, and possibly radio and optical pulsations. Since the 1970s, it has been studied at a level of detail very unusual for a neutron star. As a result, not only are its nature and distance well known, but also a good understanding has been reached of the physical processes responsible for its multiwavelength emission.

Neutron stars. Understanding neutron stars remains a fundamental challenge for modern astronomy. They represent the final state of evolution for many normal stars. Thus, their study also holds potentially great rewards for basic physics because of the special state of the matter which composes them. They lack the radiating atmosphere of stars like the Sun, and thus their only thermal emission arises from the planckian (blackbody) radiation of their surfaces. The solid crust of a neutron star, probably made out of a layer of iron nuclei, is heated by its neutron superfluid interior initially at very high temperatures by the supernova explosion that generated the star. However, the tiny dimensions of the surface (with a radius of 10 km or 6 mi), as well as its rapid cooling (down to temperatures less than 10^6 K after 10^5 years), limit the overall luminosity of the neutron star to a small fraction of that of the Sun. As a result, the classical channel of optical astronomy has been largely unable to provide information about these stars. *See* HEAT RADIATION; SUPERNOVA.

Owing to their large, rapidly spinning magnetic fields, however, neutron stars also emit significant nonthermal luminosity, that is, radiation produced by the interaction of accelerated particles and electromagnetic fields, or by the spinning fields themselves. Historically, this radiation was first detected at radio frequencies with the discovery of pulsars in 1968. Subsequently, notably with the launch of *COS-B* in 1975 and of the *Compton Gamma-Ray Observatory* in 1991, came the observation of gamma-ray pulsars. These are radio pulsars that emit a large flux of high-energy (100-MeV) gamma-ray photons. At the same time, a gamma-ray source was discovered that had all the markings of being connected with a radio pulsar but had no such object in its optical positional error box in the sky. *See* GAMMA-RAY ASTRONOMY; RADIO ASTRONOMY; SATELLITE (ASTRONOMY).

Nature and radiation. The object, named Geminga (an acronym for the constellation Gemini, and a gamma-ray source therein but also derived from a phrase in Milanese dialect meaning "it does not exist"), is indeed a rotating neutron star. It has, however, a number of unique properties: It was the first unidentified gamma-ray source in the sky, and subsequently became the first isolated neutron star to be discovered through its x-ray and gamma-ray emission, without the help of radio astronomy. High-precision positional measurements using the *Hipparcos* satellite and the *Hubble Space Telescope* allowed, for the first time, the absolute astrometry of an object this faint. Through a delicate space-time correlation, this work has made it possible to count every single rotation of the neutron star (at the rate of about 4 per second) throughout a period of over 25 years. *See* ASTROMETRY; HUBBLE SPACE TELESCOPE; X-RAY ASTRONOMY.



Superposition of x-ray images of Geminga. North is up, east is left. Shadings and contours show inner part of image obtained by the EPIC/MOS1 and MOS2 instruments on the *XMM/Newton* x-ray telescope (0.3–5.0-keV spectral range). Small black squares show image obtained by the ACIS instrument on the *Chandra X-ray Observatory* (0.45–7 keV, 0.492 arcsec pixel size), rebinned to a pixel size of 2 arcsec. The ACIS field of view is marked by the elongated white rectangle. The square box identifies a 1 × 1 arcmin area. In the *XMM/Newton* image, two elongated tails of diffuse emission are seen to emerge from the source, and show a remarkable symmetry with respect to the pulsar proper motion direction, marked by the arrow. Although insufficient in exposure and coverage to complement the *XMM/Newton* data, the *Chandra* exposure does show the existence of a thin 25-arcsec-long tail trailing the source and aligned to the object's motion, like the larger *XMM/Newton* tails. (*From A. DeLuca et al., On the complex x-ray structure tracing the motion of Geminga, Astron. Astrophys., 455:L9–L13, 2006*)

Optical and ultraviolet observations, carried out from the ground and from space, have allowed the measurement of Geminga's annual parallax, defining its distance to be about 150 parsecs (450 light-years), and thus determining with accuracy the star's absolute energy output at each wavelength. X-ray observations with the *XMM/Newton* telescope have shown that the thermal radiation from the neutron star surface has two components: a "cool" one (0.5×10^6 K) coming from a large surface and a "hot" one (10^6 K) coming from a football-field-size hot spot, probably close to the polar region and heated by high-energy particles impinging on the neutron star surface. Since both thermal emission components are modulated by the neutron star rotation, for the first time it is possible to clearly "see" the rotating surface of Geminga down to detail as minute as a 100-m^2 (1000-ft^2) portion of the surface. The particles needed to heat the hot spot are similar to those required to produce the copious gamma-ray flux from Geminga, thus providing the missing link between the x-ray and gamma-ray emission from Geminga. *See* PARALLAX (ASTRONOMY).

Wide-band data have shown that both thermal and nonthermal processes are at work in this neutron star, which is approximately 350,000 years old. Such processes originate on its surface as well as in the extremely intense stellar magnetic field. They may be another unique feature of Geminga, a very promising astrophysical laboratory.

Interaction with interstellar medium. A diffuse emission of x-rays has been discovered surrounding Geminga with the *XMM/Newton* and *Chandra* x-ray observatories. This emission has a complex, elongated shape, precisely aligned with the object's proper motion (see **illustration**). The diffuse x-rays have a hard, nonthermal spectrum, suggesting that they originate from the interaction of extremely energetic pulsar electrons with the interstellar magnetic field.

The shape of the diffuse emission traces the bow shock created by Geminga as it moves supersonically through the interstellar medium. Interestingly, the lifetime of the radiating electrons in the diffuse emission corresponds precisely to the pulsar displacement time over the size of the emission region. This agreement clinches the case both for high-energy electron acceleration by pulsars and for active pulsar interaction with their surrounding interstellar medium. *See* SHOCK WAVE.

Closer to the central pulsar, high-resolution *Chandra* data offer additional information. Thanks to an x-ray optics that allows for 10 times sharper images than the *XMM* ones, *Chandra* detects a short tail of emission trailing Geminga, again along its proper motion direction. Its origin is probably different from that of the large-scale diffuse emission seen by *XMM/Newton*. Closer to the pulsar, accelerated electrons lose their energy in the pulsar-generated magnetic field, which is much stronger than the diffuse interstellar one. With a total of 50 photons, a clear indication of a hard spectrum is present in the *Chandra* data, pointing to this pulsar-wind electron synchrotron mechanism. Unfortunately, the limited field of view of the *Chandra* observation does not cover the large-scale *XMM* tails, thus preventing a direct comparison of the results obtained by the two x-ray observatories. While *XMM/Newton* is more sensitive to large-scale features, *Chandra's* superior resolution can explore the immediate surroundings of the neutron star.

The resulting combined Geminga image represents the most complete description of a magnetized neutron star in its environment so far. Geminga is the first neutron star showing both an outer, large-scale bow shock and an inner comet-like pulsar tail. *See* CHANDRA X-RAY OBSERVATORY; NEUTRON STAR; PULSAR.

Giovanni F. Bignami

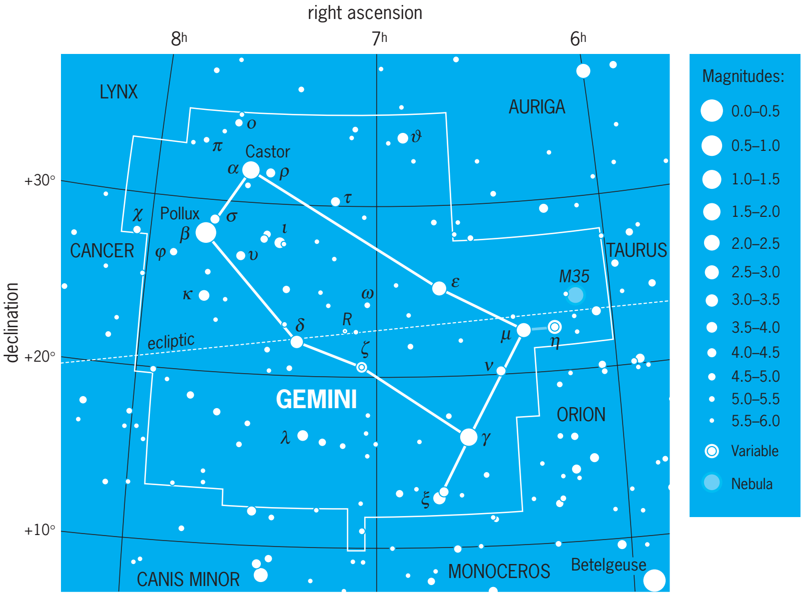
Bibliography. G. F. Bignami, Astronomy of the invisible, *Phys. World*, 16(9):27–30, September 2003 (http://physicsweb.org/articles/world/16/9/7/1); G. F. Bignami and P. A. Caraveo, Geminga: Its phenomenology, its fraternity, and its physics, *Annu. Rev. Astron. Astrophys.*, 34:331–381, 1996; P. A. Caraveo et al., Geminga's tails: A pulsar bow shock probing the interstellar medium, *Science*, 301:1345–1347, 2003; A. DeLuca et al., On the complex x-ray structure tracing the motion of Geminga, *Astron. Astrophys.*, 455:L9–L13, 2006.

Gemini

The Twins, a winter zodiacal constellation (see **illustration**). It is marked by a pair of bright stars, Castor and Pollux, usually pictured in the heads of the twins. This pair is easy to pick out with the unaided eye since they are of similar brightness and are separated by less than a fist's width when held at the end of the observer's outstretched arm. *See* ZODIAC.

In Greek mythology, Castor was a mortal, the son of the king of Sparta, and his half-brother Pollux was a son of Zeus. Both were sons of Leda, the queen of Sparta. After Castor died and went to Hades, the brothers could not bear the separation, and were reunited permanently in the heavens.

Castor is actually a multiple star containing two double stars (known to be double from their spectra) and two additional stars. Gemini includes part of the Milky Way, so it is rich in star clusters. *See* BINARY STAR; STAR CLUSTERS.



Modern boundaries of the constellation Gemini, the Twins. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Gemology

The science of those minerals and other materials which possess sufficient beauty and durability to make them desirable as gemstones. It is concerned with the identification, grading, evaluation, fashioning, and other aspects of gemstones.

For many years the little systematic attention given the subject was divided between jewelers and mineralogists. Mineralogy itself gained the status of an exact science only in the nineteenth century. Gemology was given impetus by the accelerating developments in synthetics, imitations, various means of color alteration practiced to defraud, and a growing need for developing methods of detecting these products. *See* GEM.

Richard T. Liddicoat, Jr.

Gene

The basic unit in inheritance. There is no general agreement as to the exact usage of the term, since several criteria that have been used for its definition have been shown not to be equivalent.

The nature of this difficulty will be indicated after a description of the earlier position. The facts of mendelian inheritance indicate the presence of discrete hereditary units that replicate at each cell division, producing remarkably exact copies of themselves, and that in some highly specific way determine the characteristics of the individuals that bear them. The evidence also shows that each of these units may at times mutate to give a new equally stable unit (called an allele), which has more or less similar but not identical effects on the characters of its bearers. *See* ALLELE; MENDELISM.

These hereditary units are the genes, and the criteria for the recognition that certain genes are alleles have been that they (1) arise from one another by a single mutation, (2) have similar effects on the characters of the organism, and (3) occupy the same locus in the chromosome. It has long been known that there were a few cases where these criteria did not give consistent results, but these were explained by special hypotheses in the individual cases. However, such cases have been found to be so numerous that they appear to be the rule rather than the exception. For a more complete discussion *See* MUTATION; RECOMBINATION (GENETICS).

The term gene, or cistron, may be used to indicate a unit of function. The term is used to designate an area in a chromosome made up of subunits present in an unbroken unit to give their characteristic effect. It is probable that with increasing knowledge of the nature and properties of deoxyribonucleic acid (DNA) it will become possible to reach a more generally acceptable solution to the problems of terminology. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID.

Alfred H. Sturtevant

Molecular biology. Every gene consists of a linear sequence of bases in a nucleic acid molecule. Genes are specified by the sequence of bases in DNA in prokaryotic, archaeal, and eukaryotic cells, and in DNA or ribonucleic acid (RNA) in prokaryotic or eukaryotic viruses. The flow of genetic information from DNA, to messenger RNA (mRNA), to protein is historically referred to as the central dogma of molecular biology; however, this view required modification with the discovery of retroviruses, whose genetic flow goes from RNA to DNA by reverse transcription and then to mRNA and proteins. The ultimate expressions of gene function are the formation of structural and regulatory RNA molecules and proteins. These macromolecules carry out the biochemical reactions and provide the structural elements that make up cells. *See* RETROVIRUS.

Flow of genetic information. The goal of molecular biology as it applies to genes is to understand the function, expression, and regulation of a gene in terms of its DNA or RNA sequence. The genetic information in genes that encode proteins is first transcribed from one strand of DNA into a complementary mRNA molecule by the action of the RNA polymerase enzyme. The genetic code is nearly universal for all prokaryotic, archaeal, and eukaryotic organisms. Many kinds of eukaryotic and a limited number of prokaryotic mRNA molecules are further processed by splicing, which removes intervening sequences called introns. In some eukaryotic mRNA molecules, certain bases are also changed posttranscriptionally by a process called RNA editing. The genetic code in the resulting mRNA molecules is translated into proteins with specific amino acid sequences by the action of the translation apparatus, consisting of transfer RNA (tRNA) molecules, ribosomes, and many other proteins. The genetic code in an mRNA molecule is the correspondence of three contiguous (triplet) bases, called a codon, to the common amino acids and translation stop signals (see **table**); the bases are adenine (A), uracil (U), guanine (G), and cytosine (C). There are 61 codons that specify the 20 common amino acids, and 3 codons that lead to translation stopping; hence the genetic code is degenerate, because certain amino acids are specified by more than one codon. *See* INTRON.

The sequence of amino acids in a protein is determined by the series of codons starting from a fixed translation initiation codon. AUG and GUG are the

Genetic code, showing the correspondence between triplet codons in mRNA and the 20 common (L-isomer) amino acids inserted into polypeptides by the translation apparatus

First position in codon	Second position in codon				Third position in codon
	U	C	A	G	
U	Phenylalanine	Serine	Tyrosine	Cysteine	U
	Phenylalanine	Serine	Tyrosine	Cysteine	C
	Leucine	Serine	Translation stop	Translation stop	A
	Leucine	Serine	Translation stop	Tryptophan	G
C	Leucine	Proline	Histidine	Arginine	U
	Leucine	Proline	Histidine	Arginine	C
	Leucine	Proline	Glutamine	Arginine	A
	Leucine	Proline	Glutamine	Arginine	G
A	Isoleucine	Threonine	Asparigine	Serine	U
	Isoleucine	Threonine	Asparigine	Serine	C
	Isoleucine	Threonine	Lysine	Arginine	A
	Methionine	Threonine	Lysine	Arginine	G
G	Valine	Alanine	Aspartate	Glycine	U
	Valine	Alanine	Aspartate	Glycine	C
	Valine	Alanine	Glutamate	Glycine	A
	Valine	Alanine	Glutamate	Glycine	G

major translation start codons of prokaryotic genes. AUG is almost always the translation start codon of eukaryotic genes. The bacterial start AUG and GUG codons specify a modified form of methionine, whereas AUG or GUG codons internal to reading frames specify methionine or valine, respectively. *See* GENETIC CODE.

The translation apparatus reads the next codon in the mRNA and attaches the specified amino acid onto methionine through a peptide bond. In most cases, this linear process of moving to the next codon and attaching the corresponding amino acid continues until one of the translation stop codons is encountered. Meanwhile, the nascent polypeptide chain folds by itself, or with the assistance of proteins called chaperones, into the functional protein. The biochemical rules that govern protein folding are often referred to as the second genetic code. In addition, some eukaryotic and prokaryotic proteins undergo protein splicing, which removes internal polypeptide segments called inteins. Some RNA transcripts are not translated; instead, they are cut and processed to form structural RNA molecules, such as tRNA and the three large RNA molecules associated with proteins in the ribosomes. *See* PROTEIN; RIBONUCLEIC ACID (RNA).

Isolating genes. In many cases, only genes that mediate a specific cellular or viral function are isolated. The recombinant DNA methods used to isolate a gene vary widely depending on the experimental system, and genes from RNA genomes must be converted into a corresponding DNA molecule by biochemical manipulation using the enzyme reverse transcriptase. The isolation of the gene is referred to as cloning, and allows large quantities of DNA corresponding to a gene of interest to be isolated and manipulated.

After the gene is isolated, the sequence of the nucleotide bases can be determined. The goal of the large-scale Human Genome Project is to sequence all the genes of several model organisms and humans. The sequence of the region containing the gene can reveal numerous features. If a gene is thought to encode a protein molecule, the genetic code can be applied to the sequence of bases determined from the cloned DNA. The application of the genetic code is done automatically by computer programs, which can identify the sequence of contiguous amino acids of the protein molecule encoded by the gene. If the function of a gene is unknown, comparisons of its nucleic acid or predicted amino acid sequence with the contents of huge international databases can often identify genes or proteins with analogous or related functions. These databases contain all the known sequences from many prokaryotic, archaeal, and eukaryotic organisms. Putative regulatory and transcript-processing sites can also be identified by computer. These putative sites, called consensus sequences, have been shown to play roles in the regulation and expression of groups of prokaryotic, archaeal, or eukaryotic genes. However, computer predictions are just a guide and not a substitute for analyzing expression and regulation by direct experimentation. *See* HUMAN GENOME PROJECT; MOLECULAR BIOLOGY.

Prokaryotic gene. In eubacteria, cyanobacteria (blue-green algae), and many bacteriophages, the genetic material is double-stranded DNA. However, in some bacteriophages the genetic material is single-stranded DNA or even RNA. As in other organisms, bacterial genes specify structural and regulatory RNA molecules, which do not encode proteins, or mRNA molecules, which do encode proteins. Some sites that play important cellular roles but are not copied into RNA molecules are also considered genes, such as the origin for bacterial chromosome replication. *See* BACTERIOPHAGE; CYANOBACTERIA.

Structure and expression. The arrangement of prokaryotic genes varies from simple to complex. Bacterial genes are delineated by sites in the DNA of the bacterial chromosome and in the RNA of transcripts. Transcribed genes start with promoters, which are binding sites of RNA polymerase. RNA polymerase melts the DNA near one end of each promoter region and then begins copying one of the DNA strands into an RNA molecule with a complementary sequence of bases. The starting end of bacterial RNA molecules initially contains a triphosphate group, which is not modified or capped as in eukaryotic cells. The first base is usually a purine (adenine or guanine). As RNA polymerase adds nucleotides to a growing RNA chain, it moves down the DNA molecule in one direction only until it encounters a signal to terminate the transcription process. Promoters face in both directions around the bacterial chromosome, and the orientation within promoters directs RNA polymerase one way or the other. Bacterial genes do not usually overlap extensively, and only one DNA strand is transcribed in most regions of bacterial chromosomes.

Genes that specify structural RNA molecules, such as tRNA or ribosomal RNA (rRNA) required for protein synthesis, are initially transcribed into long precursor molecules. Mature structural RNA molecules are cut from these precursor transcripts by the concerted activities of specific ribonucleases, which are enzymes that break the chemical bonds in the phosphodiester backbone of RNA. For genes that specify proteins, the nascent mRNA is translated as soon as a ribosome-binding site clears the transcribing RNA polymerase. Simultaneous synthesis and translation of mRNA molecules is a fundamental property of prokaryotic cells, because they lack the nuclear boundary that separates transcription from translation in eukaryotic cells. Segments of mRNA between the transcript beginning and the first ribosome-binding site are called leader regions and are often shorter than 100 nucleotides; however, longer leaders are found that play roles in gene regulation.

The ribosome binding site consists of two parts. The translation start codon of bacterial genes is usually AUG or GUG, which specifies a modified form of the amino acid methionine that is often removed from the final protein product. However, not every AUG or GUG in a transcript directs ribosome binding.

AUG and GUG start codons in mRNA are preceded by another short segment of nucleotides, called the Shine-Delgarno sequence. The bases in the Shine-Delgarno sequence pair with complementary bases in 16S rRNA molecules in the ribosome and properly position the start codon for translation. *See* RIBOSOMES.

Following initiation, the ribosome usually moves down the mRNA, reading one triplet (three-base) codon at a time. The amino acid corresponding to each codon in the genetic code is attached to the preceding amino acid by formation of a peptide bond. The ribosome continues synthesis of the polypeptide chain until one of the three translation stop codons is encountered, and the polypeptide and mRNA are released from the ribosome. The polypeptide chain is folded and sometimes binds to other folded polypeptides to form an enzymatically or structurally active protein. Some time after synthesizing the translation stop codon, RNA polymerase encounters a signal to stop transcription and to release the mRNA and DNA from the enzyme. Two kinds of transcription stop signals are used in bacteria. Factor-independent termination involves formation of a folded structure preceding a run of uracil residues in the nascent RNA chain. Factor-dependent termination involves interaction between a protein called Rho and RNA polymerase.

Expression level and regulation. Bacteria are ideally suited for survival as single cells in many environmental conditions. Part of this survivability involves controlling the expression of genes to optimize bacterial metabolism in response to environmental changes. For example, when a bacterium is presented with the amino acid tryptophan, a series of regulatory events are set off that turn off the genes encoding the enzymes that synthesize tryptophan. However, lack of tryptophan triggers the synthesis of these biosynthetic enzymes.

The expression level and regulation of a bacterial gene is influenced by five different processes. First, expression level depends on how frequently RNA polymerase transcribes a bacterial gene. The rate of transcription initiation depends on the relative intrinsic strength of a gene's promoter and whether transcription initiation is activated or repressed by additional protein regulatory factors that bind to DNA at or near a promoter in response to environmental factors. In addition, bacterial cells contain several kinds of RNA polymerase molecules that recognize different DNA consensus sequences as promoters in response to changing environmental conditions. Second, gene expression level depends on whether a transcribing RNA polymerase encounters a transcription termination signal called an attenuator that precedes the translated regions of certain genes. The frequency of termination at attenuator sites is sometimes controlled by environmental factors. Third, gene expression level depends on the stability of an mRNA molecule, because the longer its lifetime, the more frequently an mRNA species can be translated. Generally, bacterial mRNA molecules have chemical half-lives of only about 2 min, which is short compared to most eukaryotic mRNA. Fourth, gene expression level depends on the efficiency of translation of a given mRNA molecule. This efficiency depends on the relative intrinsic strength of a ribosome binding site and whether access of ribosomes to a binding site is regulated by protein factors or folded structures in the mRNA transcript. Finally, gene expression level depends on the relative stability of the gene product RNA or protein molecules. Stable RNA and protein molecules will accumulate in cells, compared to ones that are rapidly degraded. *See* BACTERIAL GENETICS; PROKARYOTAE. Malcolm E. Winkler

Eukaryotic gene. Eukaryotic genes are arranged in a linear array on one set of chromosomes (haploid germ cells) or two sets of chromosomes (diploid somatic cells). There are about 100,000 genes in the mammalian genome, located within chromatin at specific sites in the nucleus. Eukaryotic organelles (such as mitochondria or chloroplasts) also contain genomes that encode a much smaller number of proteins.

Transcription. In eukaryotes, transcription and translation are compartmentalized, respectively, in the nucleus and in the cytoplasm. Eukaryotic genes are transcribed by three different RNA polymerases that use the ribonucleotide triphosphates as substrates and the DNA as the template: polymerase I transcribes the larger rRNA genes; polymerase II generates mRNA by transcribing genes with open reading frames that encode proteins (the enzyme also generates certain small, nuclear RNAs that complex with nuclear proteins and play important roles in splicing); polymerase III transcribes the genes for tRNA and small rRNA as well as those for other small nuclear RNAs of mostly unknown functions. The initiation of transcription of mRNA takes place in eukaryotic genes at multiple sites, usually 30–100 base pairs downstream from a short adenine thymine–rich sequence referred to as the TATA box.

Eukaryotic genes are usually mosaics of coding (exons) and noncoding (introns) sequences. Since the pattern of removal of the introns (splicing) from the transcripts of a single gene may vary, a gene can encode several related proteins. When this happens, it is at variance with the long-standing one-gene-one-protein hypothesis. Mature, spliced mRNA (and in rare cases proteins) generates encoded proteins that lack a colinearity between the sequence of the DNA and its encoded protein. *See* EXON; TRANSPOSONS.

The polymerase II transcript is posttranscriptionally modified by capping of its 5′ end, by cleavage and polyadenylation of the 3′ end, and by splicing before the mature mRNA is transported to the cytoplasm to be translated.

Translation. Translation of eukaryotic mRNAs almost always initiates at the first ATG codon, probably controlled by ribosomes and associated initiation factors that scan the mRNA from its 5′ end. Prokaryotic protein initiation takes place at internal ATGs, depending on the location of the Shine-Delgarno sequence complementary to a rRNA sequence.

Operons have not been detected in eukaryotes. Genes that are not contiguous but coordinately regulated are members of regulons. Both eukaryotes and prokaryotes (for example, *GAL* in yeast and *arg* in *Escherichia coli*) have some of their genes organized in this manner. The expression of some genes is inducible, responding to specific changes in cell environment; others are constitutively expressed, at a uniform level. Genes that carry out specific functions may be expressed in a tissue-specific manner and regulated developmentally.

The regulation of eukaryotic genes with open reading frames is more complex than that of prokaryotic genes. Several cis-acting upstream elements (called promoters) are involved; these serve as binding sites for multiple positive and negative trans-acting gene-specific and general transcription factors. These factors, together with cofactors, interact with polymerase II to regulate its activity. Also, the regulation of some eukaryotic genes may be controlled by chromatin structure and by DNA upstream or downstream enhancer sequences that act independently of their distance (some located up to several thousand base pairs away) and/or orientation. *See* EUKARYOTAE; GENETIC CODE; GENETIC ENGINEERING; GENETICS. Julios Marmur

Bibliography. E. A. Birge, *Bacterial and Bacteriophage Genetics: An Introduction*, 4th ed., 2000; M. Karin (ed.), *Gene Expression: General and Cell-Type-Specific*, 1993; B. Lewin, *Genes V*, 1994; S. R. Maloy, J. E. Cronan, Jr., and D. Freifelder, *Microbial Genetics*, 1994; V. A. McKusick, *Mendelian Inheritance in Man: Catalogs of Autosomal Dominant, Autosomal Recessive, and X-Linked Phenotypes*, 2 vols., 12th ed., 1998; P. J. Russell, *Fundamentals of Genetics*, 1994; J. D. Watson et al., *Molecular Biology of the Gene*, 1997; E. Wingender, *Gene Regulation in Eukaryotes*, 1993.

Gene amplification

The process by which a cell specifically increases the copy number of a particular gene to a greater extent than it increases the copy number of genes composing the remainder of the genome (all the genes which make up the genetic machinery of an organism). It is therefore distinguished from duplication, which is a precise doubling of the genome preparatory to cell division, and endoreduplication, which leads to endopolyploidy.

Gene amplification results from the repeated replication of the deoxyribonucleic acid (DNA) in a limited portion of the genome, in the absence of or to a much greater extent than replication of DNA composing the remainder of the genome. Thus is formed a cell in which the genes composing a limited portion of the genome are present in relatively high copy number, while the genes composing the remainder of the genome are present in approximately normal copy number. *See* DEOXYRIBONUCLEIC ACID (DNA).

Genes coding for ribosomal RNA. Ribosomes are the site of cellular protein synthesis. They are particles which are composed of a specific type of ribonucleic acid (RNA), ribosomal RNA (rRNA), and some 30 specific proteins called ribosomal proteins. The genes coding for rRNA are multicopy genes (there are many identical or nearly identical copies in the genome of most organisms).

Gene amplification was first described in the oocytes of certain amphibians. The frog oocyte contains approximately 400,000 times as much rRNA as a normal frog liver cell in spite of the fact that the amount of DNA in the oocyte nucleus is approximately the same as that of the liver cell. It has been estimated that, by using the normal genome number of rRNA genes, it would take the frog oocyte approximately 1000 years to synthesize the amount of rRNA found in the oocyte. By amplifying these genes approximately to the level known to occur (1000- to 2000-fold), the oocyte can synthesize this amount of rRNA in 6–9 months.

The developmental significance of such amplification lies in the fact that no rRNA is synthesized during early embryogenesis in amphibians. Proteins synthesized during early embryogenesis are synthesized on maternal ribosomes that were produced by the developing oocyte. Amplification of rDNA in the amphibian oocyte occurs extrachromosomally by a rolling-circle mechanism of DNA synthesis similar to that which occurs when certain bacteriophage viruses replicate their DNA.

Amplification of rDNA also occurs in the macronucleus (the vegetative nucleus) of certain ciliated protozoans. The macronucleus grows by polyploidization, but during macronuclear growth the genes coding for rRNA are replicated to a far greater extent than is the DNA composing the remainder of the genome; that is, the rRNA genes are amplified. *See* RIBONUCLEIC ACID (RNA); RIBOSOMES.

Single-copy genes. Since gene amplification increases the copy number of a specific region of the genome without altering the copy number of genes composing the remainder of the genome, it would appear to offer an alternative method for developmental control of gene expression. By increasing the number of copies of a particular gene, the number of gene copies available for transcription could thereby be increased.

In the ovary of insects, follicle cells produce the proteinaceous egg shell (or chorion). In the ovary of the fruit fly, *Drosophila melanogaster*, the chorion is produced during the final 5 h of egg-chamber development. Production of the chorion involves the synthesis of messenger RNAs (mRNAs) for several chorion proteins within a very brief period. The genes coding for the various chorion proteins are clustered within the genome, each gene present in single-copy number in germ-line cells. Amplification results in a 15- to 50-fold increase in the number of chorion gene copies per haploid genome. Chorion gene amplification appears to enable the genes to produce the amounts of mRNA necessary to build the egg shell.

Although the mechanism of chorion gene amplification has not been completely elucidated, the data suggest that each chorion gene cluster contains a specific origin of DNA replication. Multiple rounds of replication beginning at this origin give rise to a multiforked structure in which one strand branches into two, each of these branch again and again, and the branched structures contain copies of the chorion gene.

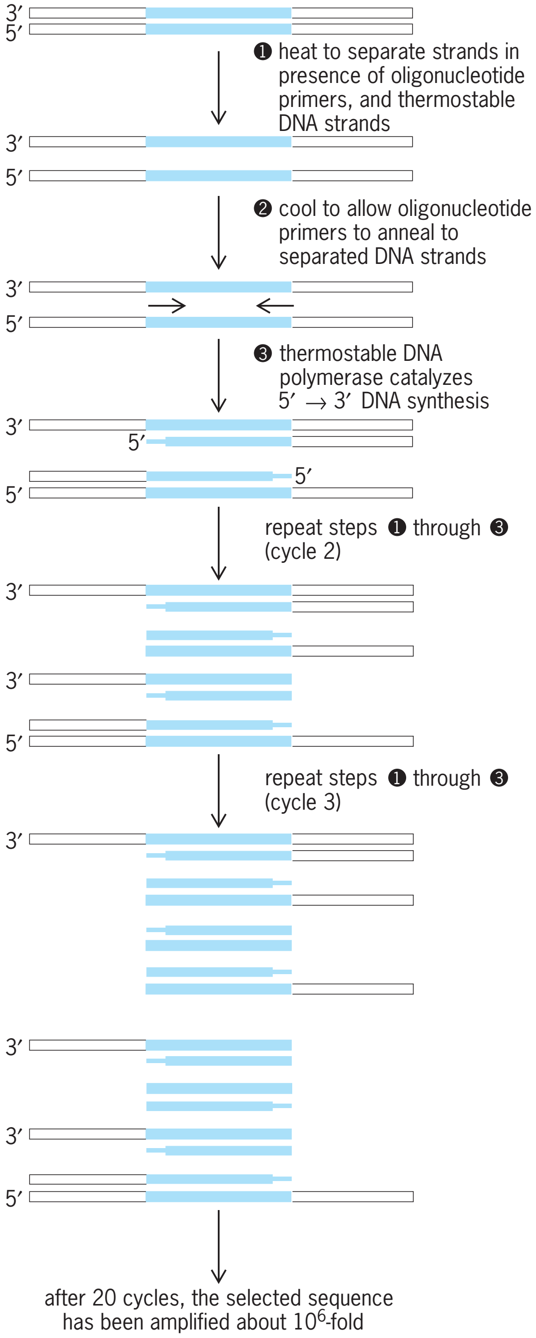
Another single-copy gene which appears to be amplified during development is one of the genes coding for the muscle protein actin during development of muscle cells (myogenesis) in the chick. The actin genes constitute a small family of genes which are present in single or very low copy numbers in the germ line. During myogenesis (formation of skeletal muscle cells), the actin genes coding for skeletal muscle cell actin are amplified approximately 85-fold.

Of course, not all single-copy genes which are responsible for the synthesis of a large amount of cell product at a particular stage of development are amplified. The silk gland of the larval silk moth, *Bombyx mori*, for instance, produces copious amounts of silk fibroin protein. Analysis of DNA from silk glands indicates that the silk fibroin genes are not amplified above the level of ploidy expressed by the silk glands. Similarly, amplification of globin genes is not a factor in the ability of the reticulocyte (precursor cell of the red blood cell) to synthesize large amounts of the protein hemoglobin.

Other examples. In the above-mentioned examples of gene amplification, the amplification phenomenon appears to be developmentally regulated, and the amplified copies of the gene are subsequently lost from the cell. Studies on cells in culture have demonstrated “amplification” of genes involved in resistance to specific drugs. Resistance to the chemotherapeutic drug methotrexate by cultured mouse cells is associated with amplification of the dihydrofolate reductase gene. In this case the amplified genes are passed to daughter cells at the time of cell division. Such gene amplification can be either retained by cells or lost from cells when they are grown in the absence of methotrexate. *See* DRUG RESISTANCE.

Polymerase chain reaction. The polymerase chain reaction (PCR) is a technique that amplifies DNA sequences in laboratory cultures. Beginning with a sample that may contain only one DNA molecule, a selected sequence within that DNA molecule can be amplified millions or even billions of times.

To amplify a given segment of DNA, its entire sequence need not be known, but only short sequences near the ends of the segment. Short oligonucleotides, complementary to the end sequences and designed so that their 3′ ends orient toward the interior of the segment, are synthesized by conventional chemical means. The DNA in the sample is denatured by heating and then is cooled in the presence of the oligonucleotides, which anneal to the separated DNA strands. The annealed oligonucleotides are then utilized by DNA polymerase as primers for DNA synthesis, replicating the target DNA segment. These simple steps are repeated over and over again, doubling the amount of the target segment in each cycle (see **illus.**). The use of thermostable DNA polymerase (isolated from bacteria that live naturally in hot springs) is an important refinement. Since the polymerase is not denatured when the DNA is heated, it need not be added at every cycle. Amplification of DNA by the polymerase chain reaction is carried out conveniently by inexpensive machines that produce the amplified product DNA in a few hours. Target sequences between a few hundred and 10,000 base pairs in length are readily amplified. After 20 cycles, the target DNA is amplified more than a millionfold; after 30 cycles the amplification is over a billionfold.



Amplification of DNA using the polymerase chain reaction. The colored region is the target DNA to be amplified. The shorter arrows represent the oligonucleotide primers.

The utility of polymerase chain reaction can be illustrated with a few examples. (1) For molecular biologists, this technique simplified the cloning of known genes, and led to the development of powerful methods to alter DNA sequences in laboratory cultures and detect alterations in DNA in organisms. (2) Certain parts of the human genome are sufficiently variable that they can be used to identify a particular individual unambiguously. With the aid of the polymerase chain reaction, DNA from samples as small as a single human hair can be amplified and analyzed to place an individual at the scene of a crime. Similar techniques are used to accurately establish paternity. Human DNA extracted frm centuries-old burial sites and amplified by this technique is being used to trace prehistoric human migrations. (3) As the molecular genetic basis for more genetic diseases becomes known, polymerase chain reaction–based diagnostic tests for these diseases are possible. For example, prenatal diagnosis of genetic diseases can be carried out with minute samples of fetal tissue. (4) Polymerase chain reaction has also been used to amplify and study DNA from extinct species such as the quagga and the woolly mammoth. In addition, very small amounts of DNA added to the cargo hold of an oil tanker can be used to trace the origins of an oil spill. *See* FORENSIC MEDICINE; GENE; GENETIC ENGINEERING; MOLECULAR BIOLOGY; PRENATAL DIAGNOSIS.

Michael M. Cox

Bibliography. W. Bloch, A biochemical perspective of the polymerase chain reaction, *Biochemistry*, 30(11):2635–2647, 1991; H. A. Erlich and N. Arnheim, Genetic analysis using the polymerase chain reaction, *Annu. Rev. Genet.*, 26:479–506, 1992; S. Paabo, R. G. Higuchi, and A. C. Wilson, Ancient DNA and the polymerase chain reaction: The emerging field of molecular archaeology, *J. Biol. Chem.*, 264(17):9709–9712, 1989; R. Schimke (ed.), *Gene Amplification*, Cold Spring Harbor Laboratory, 1982; A. C. Spradling and A. P. Mahowald, Amplification of genes for chorion proteins during oogenesis in *Drosophila melanogaster*, *Proc. Nat. Acad. Sci. U.S.A.*, 77:1096–1100, 1980; R. J. Steffan and R. M. Atlas, Polymerase chain reaction: Applications in environmental microbiology, *Annu. Rev. Microbiol.*, 45:137–161, 1991.

General aviation

All aircraft activity not associated with major airlines or the military. Among all classifications of aviation in the United States, general aviation consists of the largest number of aircraft and pilots and accounts for the largest number of flying hours.

General aviation is an integral part of the transportation system of the United States, serving both business and pleasure travel. Business travel, which accounts for more flying hours than any other branch of general aviation, employs aircraft ranging in size from single-engine vehicles flown by one person to jet equipment with a professional crew of two. Approximately 70% of all general aviation is associated with some commercial activity, such as business travel, construction, aerial application of fertilizer and pesticides, or flight instruction. Except for business jets, the average general aviation aircraft is much slower than the typical airliner, but the estimated 4.28×10^9 mi (6.89×10^9 km) flown by general aviation in 1986 exceeded by over 1×10^8 mi (1.6×10^8 km) the distance flown by scheduled carriers. General aviation departures numbered 4.76×10^7 during 1986, more than five times the number for the major airlines.

Aircraft. A general aviation aircraft can range in size from the smallest pleasure craft designed and built by an individual to the largest commercial air transport. Surplus military aircraft owned by individuals or businesses are also included in the category of general aviation, a classification based on use rather than size or original purpose.

Of the more than 220,000 active general aviation aircraft registered by the Federal Aviation Administration (FAA), almost 80% are single-engine vehicles powered by intermittent-combustion engines. Although many of these aircraft carry fewer than three passengers at speeds below 150 mi/h (240 km/h), most can carry four to six passengers at speeds up to 200 mi/h (320 km/h). Approximately 11% of the general aviation fleet are twin-engine aircraft, powered by intermittent-combustion engines and capable of cruising at speeds of 180–250 mi/h (290–400 km/h) with six to ten passengers. *See* AIRPLANE; RECIPROCATING AIRCRAFT ENGINE.

In addition, approximately 6000 multiengine aircraft are powered by turboprop engines and 4500 by turbojet or turbofan power plants. Jet aircraft employed within general aviation usually carry from five to fifteen passengers, depending on the model, with a crew of two pilots. Turboprop aircraft may also be flown by two pilots, although this is not required for all models; passenger capacity typically ranges from six to nine. Cruising speeds for a turboprop aircraft are about 300 mi/h (480 km/h), but some can exceed 400 mi/h (640 km/h) and one model can reach 460 mi/h (740 km/h) in level flight. Turboprop aircraft usually fly at altitudes between 20,000 and 40,000 ft (6000 and 12,000 m). Jet aircraft designed for general aviation typically cruise between 430 and 520 mi/h (690 and 835 km/h) and at altitudes up to 45,000 ft (13,700 m), but some are capable of cruising at speeds of 560 mi/h (900 km/h) at 51,000 ft (15,500 m), which is above the altitude approved for all commercial airliners except the Concorde. *See* JET PROPULSION; TURBOFAN; TURBOJET; TURBOPROP.

Rotorcraft number about 7000, of which approximately 4000 are powered by turboshaft engines and the rest by intermittent-combustion engines. There are also approximately 7000 gliders in the general aviation fleet. *See* GLIDER; HELICOPTER.

Certification standards. Nearly all general aviation craft weighing more than 12,499 lb (5669 kg) are certified by the federal government under Federal Air Regulation (FAR) Part 25, the same standard applied to the large commercial transports flown by major airlines. Lighter airplanes are certified under

FAR Part 23, which is only slightly less stringent in its safety specifications. Helicopters must meet the requirements of FAR Part 27 or FAR Part 29, depending upon their weight and purpose.

Pilots. General aviation pilots may hold any federal license, such as airline transport pilot, commercial, private, student, or certified flight instructor. A license, however, is valid only in conjunction with a current medical certificate of the appropriate class. At the minimum a student license is needed to pilot an aircraft, but to carry passengers a more advanced license is required, together with the successful completion every 24 months of a biennial flight review (BFR) of the pilot's knowledge and skill.

In addition to a license, current medical certificate, and biennial flight review, pilots must possess a rating for the category and class of vehicle they wish to fly. Category refers to a broad classification of aircraft, such as airplane, rotocraft, glider, or lighter-than-air balloon. Class refers to the classification of aircraft within a category having similar operating characteristics, such as single-engine, multiengine, land, sea, helicopter, gyroplane, airship, or free balloon. To operate aircraft weighing more than 12,499 lb (5669 kg) and for all aircraft powered by a turbojet or fan-jet engine, pilots must posses a specific type rating affixed to their license.

To fly an aircraft during poor weather conditions, nominally defined as visibility less than 3.0 mi (5.0 km) and a cloud ceiling of less than 1000 ft (305 m), a pilot must possess an instrument rating or an airline transport pilot license; one of these is held by most general aviation pilots who fly professionally or for business purposes. *See* AIR-TRAFFIC CONTROL.

Limitations of scheduled airlines. The demand for business transportation by general aviation was stimulated by the oil embargoes of 1973 and 1979 and by the airline deregulation of October 1978. Fuel prices increased nearly 1000% between 1974 and 1981, and even the most fuel-efficient airliners are now cost-effective only when flown with high load factors, that is, with most seats occupied by paying passengers.

Deregulation permitted airlines to concentrate on heavily traveled routes such as those between major cities; locations where the demand for travel was low were abandoned. Accordingly the major airlines developed a system of hub cities, where flights would connect either with small commuter lines that brought passengers from outlaying locations or with larger airliners flying between major hubs. This system became known as the hub-and-spoke system because the passenger traffic appeared to flow into the hub city along spokes of a wheel, with few if any flights between points on the circumference of the wheel. For example, flying by scheduled airlines from Wichita, Kansas, to Rochester, Minnesota, required intermediate stops at Memphis, Tennessee, and Minneapolis, Minnesota. The return flight stopped in Waterloo, Iowa, and St. Louis, Missouri. Such a trip would last nearly 14 h, but a direct flight by general aviation would take only 3.5–5 h. *See* AIR TRANSPORTATION.

Business flying. Because of the characteristics of hub-and-spoke systems, business people have increased their demand for general aviation, which also offers more control over travel arrangements. Although the major airlines provide scheduled service to approximately 394 locations and unscheduled flights to an additional 312 airports, over two-thirds of all flights were between only 30 cities. This service is significantly supplemented by general aviation, which utilizes the nearly 5800 public-use landing places and many of the approximately 12,000 limited-use landing places in the United States.

With a small single-engine aircraft accommodating four or five passengers, a business person with a private license and perhaps an instrument rating can escape many of the limitations of airline schedules based on the hub-and-spoke system. Some pilots prefer small or medium-size twin-engine aircraft for business travel, particularly in night and instrument flying. Some companies own their aircraft and employ a professional crew; others prefer chartering from an aircraft taxi or management firm. Business jets that fly nonstop more than 5000 mi (8100 km) are operated by many multinational corporations, offering security as well as efficiency.

Fuel consumption. General aviation aircraft are relatively fuel-efficient. A six-passenger single-engine airplane can deliver more than 80 passenger miles per gallon of fuel (34 km/liter); typical cabin-class twin-engine aircraft generate about 50 passenger miles per gallon (21 km/liter). General aviation consumes less than 0.5% of the gasoline used in the United States and 6.5% of the turbine fuel, accounting for about 0.75% of the output of United States oil refineries and representing an efficient use of energy for transportation. *See* AIRCRAFT FUEL.

Commercial applications. Commuter and regional airlines supplement the hub-and-spoke scheduling typical of commercial air travel since deregulation. Unlike business flying, which operates under FAR Part 91, commercial activities within general aviation fall under FAR Part 135. All airlines flying aircraft with more than 30 passenger seats operate under FAR Part 121, which covers both air taxis and scheduled commuter and regional airlines. Commuter and regional airlines offer transportation over prescribed routes as specified in a published schedule, whereas air taxis fly passengers and freight on demand over random routes.

Commuter and regional airlines. With the growth of hub-and-spoke air travel, small airlines offer the only scheduled service in some communities, and often cooperate with major airlines by sharing code designations used in computer reservation systems and by travel agencies. As a result of code-sharing, which became widespread in 1984, travelers often book tickets on commuter or regional carriers without realizing it. These carriers have become more sophisticated and meet many of the same standards as the major airlines.

Commuter and regional airlines might be compared with the local service airlines in the period before deregulation, which operated under FAR Part

121 with equipment smaller than that used by the major carriers.

However, since the fuel crises of 1973 and 1979 and airline deregulation in 1978, commuter and regional air carriers have experienced the same problems as the major airlines. They must choose routes that generate high load factors. Thus, they have concentrated on hub locations and moved into routes previously flown only by the major carriers. Consequently, aircraft operated by companies or individuals often offer the quickest and most cost-effective transportation between locations that are outside the routes connecting major hubs.

As a result of these changes, there is a question as to whether commuter and regional airlines should still be classed with the general aviation, or whether they should have a category of their own, or whether they really belong with the major airlines. The same FAA field organization that monitors commuter and regional operations also audits the major airlines, and the National Transportation Safety Board now places commuter and regional airline accidents within a category separate from the commercial segment of general aviation. In the United States, however, the same manufacturing, maintenance, and personnel infrastructure supports general aviation and the commuter and regional airlines.

Air taxis. Air taxis provide on-demand air transportation for hire in a wide variety of aircraft, operating under FAR Part 135. They do not fly on a schedule over established routes. The demand for increased service is being met by operators ranging from small flying services with one four-passenger single-engine airplane for occasional charter work to large transportation companies with a fleet of business jets. Under FAR Part 91, the air taxi segment of general aviation manages and operates aircraft owned by a business or shared by more than one business.

Agriculture. General aviation is widely used in agriculture; aerial seeding, fertilizing, and spraying are efficient and widely used by farmers. Over 1.8×10^8 acres (7.3×10^7 hectares) of farmland are treated in the United States each year in this way. For example, about 90% of the United States rice crop is seeded by air. A day's work with a tractor is accomplished by an aircraft in 1 h with a 90% fuel saving. Chemicals are applied more effectively by air than by tractor. *See* AGRICULTURAL AIRCRAFT.

Flight instruction. Flight instruction represents about 15% of all general aviation activity and is the principal source of professional pilots. Most flight instruction and air taxi services are provided by fixed-base operators at an airport. A typical establishment rents aircraft and offers instruction as well as taxi or charter flying. It may also sell and service new or used equipment; it usually also sells gas and oil. The number of fixed-base operator chains and franchises has increased as the service segment of general aviation has matured.

Private flying. About 30% of general aviation involves the personal use of small aircraft for transportation and leisure purposes as well as some of the proficiency flying required by the Federal Aviation Administration of pilots who wish to maintain their licenses. Most personal flying employs small single-engine craft.

The use of gliders, sports aircraft, balloons, and home-built vehicles is considered personal unless it involves flight instruction for profit. Home-built aircraft cannot be used for any commmercial purpose and can serve only in personal flying. *See* ULTRALIGHT AIRCRAFT.

Research applications. General aviation has often been the testing ground for innovation. Most of the pioneer aviators flew equipment that could be considered the general aviation vehicles of their day. General aviation manufacturers produced the first nonmilitary aircraft to be built entirely from composite materials and approved for routine use.

Voyager. Over a 6-year period, a team of individuals led by pilots D. Rutan and J. Yeager and designer B. Rutan conceived and implemented a plan to circumnavigate the globe by air without refueling, a feat that had never been accomplished. Their aircraft, *Voyager* (**Fig. 1**), was designed and constructed for the mission. With wings longer than a Boeing 727 airliner's, *Voyager* carried nearly 9000 lb (4000 kg) of fuel in a structure that was built of advanced composite materials and weighed less than 1000 lb (450 kg) when empty. In effect, the twin-engine aircraft was a flying fuel tank with minimal space for pilot and copilot. Its fuselage, wings, and twin booms contained aviation gasoline for a liquid-cooled engine that was located in the aft fuselage and used for cruising flight and for a larger, air-cooled engine that was located in the aircraft's nose and used in combination with the rear engine for takeoff. Once in cruising flight, the forward engine was stopped and its propeller positioned in a feathered or knife-edge orientation to minimize wind resistance.

Voyager took off from Edwards Air Force Base, California, on December 14, 1986, and landed there 9 days later without refueling. In spite of challenging weather and several mechanical failures, the crew completed their global circumnavigation without mishap or physical injury.

Human-powered flight. The dream of flight using only the aviator's physical strength was unobtainable until advanced materials and computer-aided design

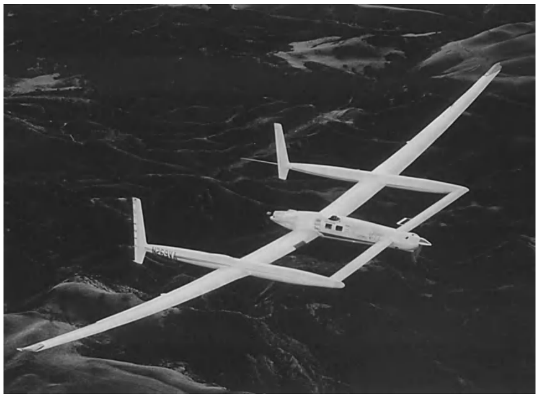


Fig. 1. ***Voyager*** **after takeoff on world flight. (*****Jeffrey Vock/Visions*****)**

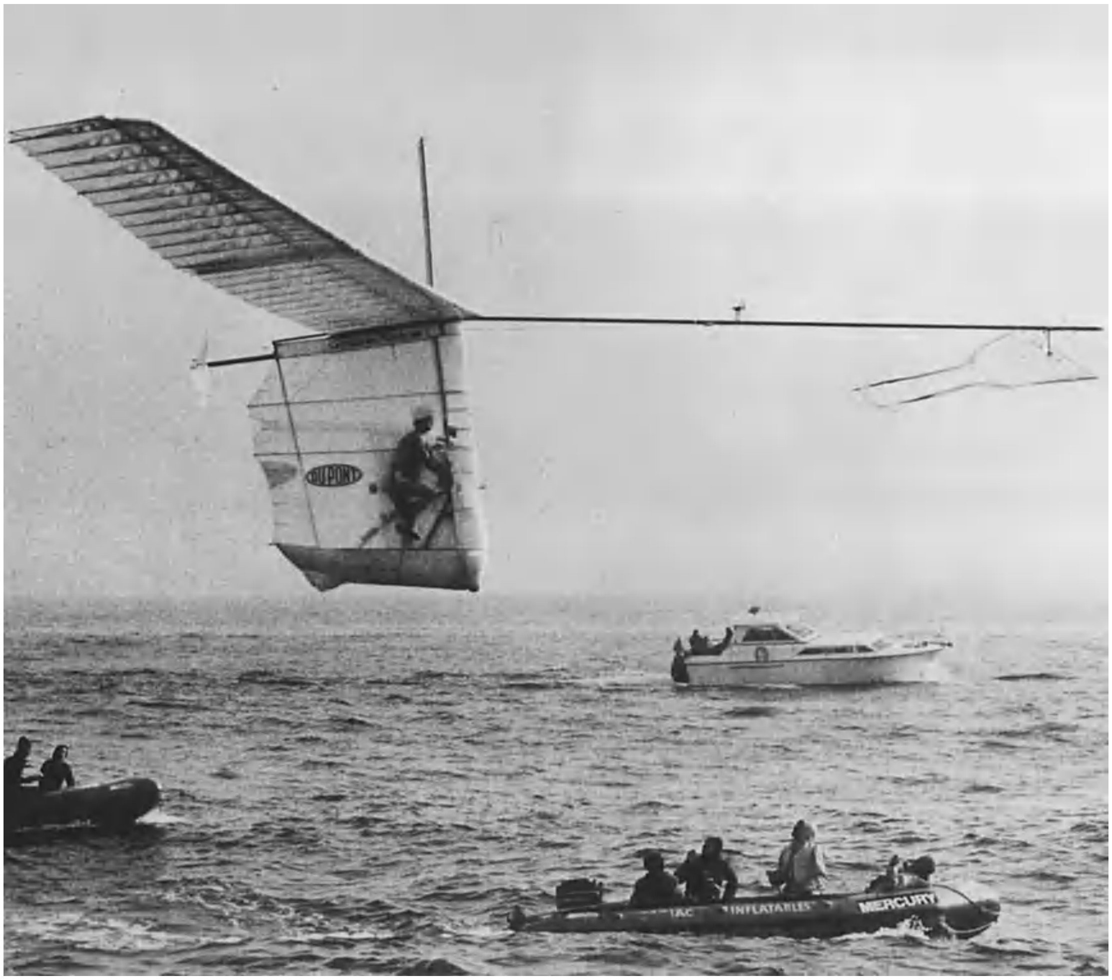


Fig. 2. Human-powered aircraft *Gossamer Albatross* in flight. (*Photograph by Don Monroe*)

provided the tools needed to design aircraft with sufficient strength-to-weight characteristics and efficiency to be powered entirely by a human. *See* COMPOSITE MATERIAL; COMPUTER-AIDED DESIGN AND MANUFACTURING.

Modern attempts to achieve human-powered flight were stimulated in 1959 when H. Kremer posted a prize for the first British aircraft to fly solely by means of human power, without the aid of power storage or buoyancy, around a figure-eight course. (The prize was later opened to all nationalities.) The aircraft had to be powered and controlled by a crew throughout its flight, and only one crewmember could assist during the takeoff. Still air, defined as a wind not to exceed 10 knots (5.1 m/s), was required for the flight in order to claim the prize.

The first practical human-powered aircraft flew in 1961 but without sufficient performance to win the Kremer prize, which went unclaimed until 1977 when it was won by the *Gossamer Condor* aircraft, which was designed by P. D. MacCready, Jr., and flown by B. Allen. The team of MacCready and Allen subsequently accomplished many significant events in human-powered flight, including the 1979 flight of the *Gossamer Albatross* (**Fig. 2**) across the English Channel.

Safety. General aviation aircraft flown in corporate operations by salaried crews have a safety record rivaling that of the major airlines. For all general aviation, including pleasure flying, the accident rate was 7.95 per 100,000 flight hours, and the fatal accident rate was 1.49 per 100,000 flight hours, in 1988. Safety has been improving steadily in general aviation. *See* AVIATION.

John W. Olcott

Bibliography. General Aviation Manufacturers Association, *General Aviation*, 1989; General Aviation Manufacturers Association, *General Aviation Statistical Databook*, 1988; *National Transportation Safety Board Annual Review of Aircraft Accident Rates*, annually; U.S. Department of Transportation, *Census of U.S. Civil Aircraft*, annually; U.S. Department of Transportation, Federal Aviation Administration, *FAA Statistical Handbook of Aviation*, annually.

Generator

A machine in which mechanical energy is converted to electrical energy. Generators are made in a wide range of sizes, from very small machines with a few watts of power output to very large central-station generators providing 1000 MW or more. All electrical generators utilize a magnetic field to produce an output voltage which drives the current to the load. The electric current and magnetic field also interact to produce a mechanical torque opposing the motion supplied by the prime mover. The mechanical power input is equal to the electric power output plus the electrical and mechanical losses.

Generators can be divided into two groups, alternating current (ac) and direct current (dc). Each group can be subdivided into machines that use permanent magnets to produce the magnetic field (PM machines) and those using field windings. A further subdivision relates to the type of prime mover and the generator speed. Large generators are often driven by steam or hydraulic turbines, by diesel engines, and sometimes by electric motors. Generator speeds vary from several thousand rotations per minute for steam turbines to very low speeds for hydraulic or wind turbines. *See* DIESEL ENGINE; HYDRAULIC TURBINE; MOTOR; PRIME MOVER; STEAM TURBINE; WIND POWER.

Theory of operation. The production of the output voltage is based upon Faraday's law, which states that when the magnetic flux across a coil of wire changes, the resulting induced voltage is proportional to the number of turns in the coil multiplied by the rate of change of the flux, as stated in Eq. (1), where e is

$$e = N\frac{d\phi}{dt} \tag{1}$$

the instantaneous induced voltage, N the number of turns in the coil, ϕ the flux in webers, and t the time in seconds. *See* ELECTROMAGNETIC INDUCTION.

Typically, voltage is induced in the armature windings of a generator by mechanically driving one member relative to the other so that the active coils (the armature winding) move through the magnetic field. In this case the flux linking the armature coils changes because of the relative motion, and the resulting voltage is called a motional or speed voltage. When the magnetic field is constant, as is usual, the motional voltage can be expressed by Eq. (2), where B is the

$$e = 2NBlv \tag{2}$$

magnetic flux density in teslas, l the length of armature conductors in meters, and v the relative velocity of the armature conductors with respect to the magnetic field in meters per second. This form of Faraday's law is applicable to most conventional generators.

When the generator supplies an electrical load, an electric current exists in the armature windings. This current interacts with the magnetic field to produce a mechanical force f which is in oppositon to the relative motion and proportional to the current and the magnetic field, as shown by Eq. (3), where i is

$$f = 2NBli \tag{3}$$

the current in the armature conductors. The power balance in the generator can be demonstrated by observing that the electrical power given by the product of voltageand current in Eq. (4) is equal to the

$$ei = 2NBlvi \tag{4}$$

mechanical power expressed by the product of force and velocity in Eq. (5).

$$fv = 2NBliv \tag{5}$$

See MAGNETIC FIELD; POWER.

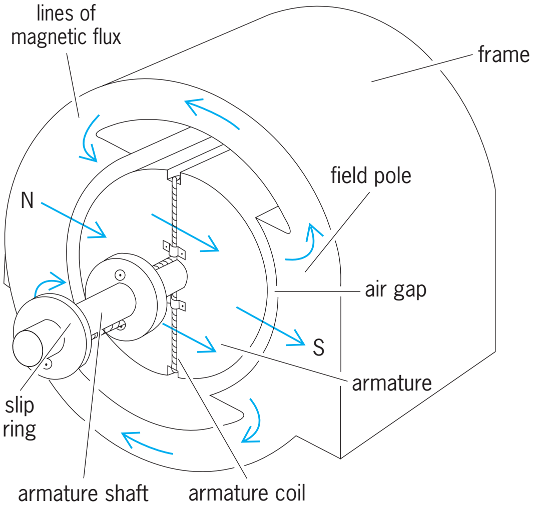


Fig. 1. Elementary generator.

Basic design. **Figure 1** shows an elementary generator with a stationary field and a single rotating armature coil. If the field is uniformly distributed along the field pole, the voltage generated as the coil moves through the field is nearly constant over each half revolution according to Eq. (2). The polarity of the voltage reverses once per revolution as the coil sides come under the influence of first a north pole and then a south pole. The resulting rectangular-shaped voltage wave has the appearance of the wave shown in **Fig. 2*a***. If the slip rings shown in Fig. 1 are replaced by a two-segment commutator, the polarity of the output can be maintained in one direction to produce the pulsating dc voltage shown in Fig. 2*b*. *See* COMMUTATOR; DIRECT-CURRENT GENERATOR.

To produce a sinusoidal shaped ac output voltage, the field poles can be modified to produce a flux distribution in which the field strength builds up gradually to a peak at the center of the pole. If the shape of the field distribution were made sinusoidal, the resulting voltage wave would, according to Eq. (2), vary sinusoidally (Fig. 2*c*). This voltage can be extracted via the slip rings shown in Fig. 1. With additional armature coils and slip rings, polyphase voltages can be produced. *See* ALTERNATING-CURRENT GENERATOR.

Although Fig. 2 illustrates a machine with only two magnetic poles, machines with a larger, even number of poles are also possible. Since one electrical cycle is generated each time an armature coil passes two magnetic poles, the rotational speed necessary

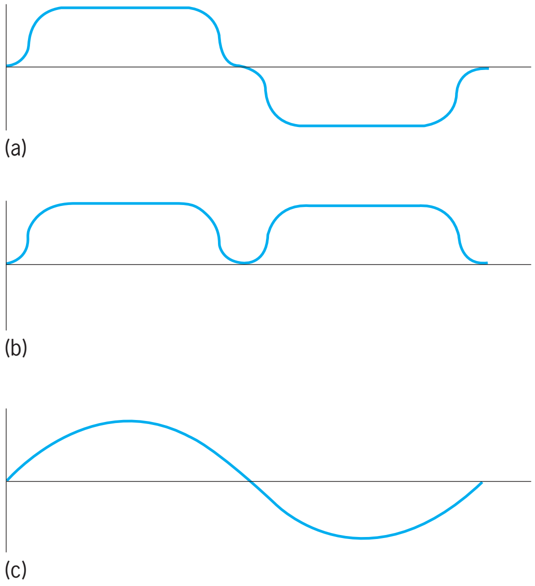


Fig. 2. Output voltage waveforms of elementary generators. (*a*) At slip-ring terminals, with uniform flux distribution. (*b*) At commutator terminals, with uniform flux distribution. (*c*) At slip-ring terminals, with sinusoidal flux distribution.

to produce a particular output frequency is given by Eq. (6), where RPM is the rotational speed in revolutions per minute (rpm), f is the frequency of the generated voltage in hertz (Hz), and P is the number of poles. An ac generator producing 60 Hz can have speeds of 3600 rpm for two poles, 1800 rpm for four poles, 1200 rpm for six poles, and so forth.

$$\mathrm{RPM} = \frac{120f}{P} \tag{6}$$

Losses, rating, and efficiency. The conversion of energy from one form to another is never a process without loss. The principal losses in generators are the mechanical losses associated with bearing friction and windage, the electrical losses resulting from the resistance of the conductors (proportional to the current squared), and the magnetic losses in the iron cores of the armature and field structures. These losses all produce heat. The allowable temperature rise of the generator system limits the total losses that can be tolerated. This, in turn, determines the rating of the generator in volt-amperes for ac generators or watts for dc generators. The rating denotes the electrical load that the generator can supply continuously without damage caused by overheating. *See* CORE LOSS; ELECTRICAL RESISTANCE; FRICTION.

The efficiency of a generator is the ratio of the electrical power output to the mechanical power input. Without power losses, the efficiency would be 100%. Large generators can reach efficiencies as high as 97% or more, but small generators are less efficient. A generator of a few hundred watts rating, such as that used to charge an automobile battery, may have an efficiency of only 60–70%. *See* EFFICIENCY.

Construction. The field structure of a generator establishes the magnetic flux needed for energy conversion. In small generators, permanent magnets can be used to provide the required magnetic field. In large machines, dc field windings are more economical and permit changes in the magnetic flux and output voltage. This allows control of the generated voltage, which is important in many applications. Electromechanical energy conversion does not take place in the field windings; the input electrical energy is all converted to heat.

In dc generators the field structure must be stationary to permit a rotating mounting for the commutator and armature windings. However, since the field windings require low voltage and power and have only two lead wires, it is convenient to place the field on the rotating member in ac generators. The field winding is more easily insulated and supported against rotational forces and is better suited to sliding contact slip rings than the higher-voltage, higher-power armature windings. Usually the field windings are mounted to encircle the protruding field poles, creating a salient pole generator (Fig. 1), but in some cases, especially for high-speed ac generators, the field windings are placed in slots in the surface of a cylindrical field structure. In larger generators a second, smaller generator called an exciter supplies the dc power required to energize the field windings. In most modern large generators the exciter is an ac generator whose output is electronically converted to dc. If the electronic rectifier components rotate with the exciter and field windings, no slip rings are needed and the combination is called a brushless exciter.

The armature and field cores are constructed of steel to minimize the field current (or size of permanent magnets) needed to establish the magnetic flux. The air gap between the cores is kept small since it is difficult to magnetize compared to the steel cores. Those portions of the core in which the magnetic flux is constant can be solid steel. This includes the field poles and outer yoke of dc generators and most of the field structure of ac generators. The armature core, which is subject to time-varying flux, is always constructed from thin sheets of high-grade electrical steel to reduce eddy-current losses. *See* EDDY CURRENT; MAGNET.

The armatures of all generators are cylindrical and do not have the protruding poles that are common for the field structure. The windings are placed in slots along the periphery of the armature structure and are electrically insulated from the core. In ac generators these slots are found along the inner surface of the stationary armature core. In dc generators the slots are found on the outer surface of the rotating armature core (Fig. 1). The slots have two important functions. First, the teeth between them provide a path for the flux. Second, most of the mechanical force which would be exerted on the windings if they were placed in the air gap transfers to the core itself. There is thus no need to support the windings to transmit the forces required in large generators. *See* WINDINGS IN ELECTRIC MACHINERY.

Power electronics in generators. The availability of power electronic converters capable of modifying the voltage and frequency of large amounts of electric power has resulted in the appearance of a number of generator systems offering reduced cost, unique properties, or both. An example, involving millions of units, is the battery-charging automotive alternator, consisting of an ac generator and a rectifier. In this case the rectifier serves to replace the commutator of the dc generator used in earlier times. In addition to lower cost, the system offers a wider useful speed range and simpler control. On a totally different scale, most modern wind turbine generators employ an inverter to adjust the frequency to a constant 60 Hz while allowing the wind turbine speed to vary to optimize the turbine efficiency. Although small in number, the largest of these generators are in the 1–5-MW size range. Many other examples of power electronics applied to generator systems exist, and their number is expected to grow. *See* ELECTRIC POWER GENERATION; ELECTRIC ROTATING MACHINERY. Donald W. Novotny

Bibliography. A. R. Bergen and V. Vittal, *Power Systems Analysis*, 2d ed., 2000; J. J. Cathey, *Electric Machines: Analysis and Design*, 2001; S. J. Chapman, *Electric Machinery Fundamentals*, 4th ed., 2005; A. E. Fitzgerald, C. Kingsley, Jr., and S. D. Umans, *Electric Machinery*, 6th ed., 2003.

Genetic algorithms

Search procedures based on the mechanics of natural selection and genetics. Such procedures are known also as evolution strategies, evolutionary programming, genetic programming, and evolutionary computation. Originating in the cybernetics movement of the 1940s and the 1950s, genetic algorithms are increasingly solving difficult search, optimization, and machine-learning problems that have previously resisted automated solution.

Motivation. Just as natural selection and genetics have filled a variety of niches by creating genotypes (attributes of chromosomes) that result in well-adapted phenotypes (attributes of organisms), so too can genetic algorithms solve many artificial problems by creating strings (artificial chromosomes) that result in better solutions. Users ultimately turn to genetic algorithms for robustness, that is, for algorithms that are broadly applicable, relatively quick, and sufficiently reliable. This emphasis on robustness contrasts starkly with the philosophy of operations research, where new algorithms must be tailored to specific problems. The need to invent a new method for each new problem class is daunting, and users look for methods that can solve complex problems without this requirement. *See* GENETICS; OPERATIONS RESEARCH; ORGANIC EVOLUTION.

Mechanics. There are many variations on genetic and evolutionary algorithms. For concrete exposition, the discussion is limited to a simple genetic algorithm that processes a finite population of fixed-length, binary strings. In practice, bit codes, *k*-ary codes, real (floating-point) codes, permutation (order) codes, program codes, and others have all been used with success.

A simple genetic algorithm consists of three operators: selection, crossover, and mutation.

Selection. Selection is the survival of the fittest within the genetic algorithm. The key notion is to give preference to better individuals. For selection to function, there must be some way of determining what is good. This evaluation can come from a formal objective function, it can come from the subjective judgment of a human observer or critic, or it can come from the co-evolution of one's competitors or collaborators. There are many ways to achieve effective selection, including ranking, tournament, and proportionate schemes. For example, in two-party tournament selection, pairs of strings are drawn randomly from the parental population, and the better individual places an identical copy in the mating pool. If a whole population is selected in this manner, each individual will participate in two tournaments, and the best individual in the population will win both trials, the median individual will typically win one trial, and the worst individual wins not at all. As this example makes clear, the primary requirement of selection is for a partial ordering.

Crossover. If genetic algorithms were to do nothing but selection, the trajectory of populations could contain nothing but changing proportions of the strings in the original population. To do something more sensible, the algorithm needs to explore different structures. A primary exploration operator used in many genetic algorithms is crossover. Simple, one-point crossover begins when two individuals are chosen from the population by using the selection operator, and these two structures are considered to be mated. A cross site along the string length is chosen uniformly at random, and position values are exchanged between the two strings following the cross site. For example, starting with the two strings $A = 11111$ and $B = 00000$, if the random choice of a cross site turns up a 3, the two new strings $A' = 11100$ and $B' = 00011$ are obtained following crossover; these strings would be placed in the new population. This process continues pair by pair until the new population is complete, filled with so-called offstrings that are constructed from the bits and pieces of good (selected) parents. There are many other variants of crossover, but the main issue is whether the operator promotes the successful exchange of necessary substructures.

Selection and crossover are surprisingly simple operators, involving nothing more complex than random-number generation, string copying, and partial string exchanges, yet their combined action is responsible for much of a genetic algorithm's effectiveness. To understand this statement intuitively, it is useful to think in terms of human process of innovation. Innovation or creativity is often based on combining notions that work well in one context with notions that work well in another context to form new, possibly better ideas of how to attack the problem at hand. Similarly, genetic algorithms juxtapose many different, highly fit substrings (notions) through the combined action of selection and crossover to form new strings (ideas).

Mutation. In a binary-coded genetic algorithm, mutation is the occasional (low-probability) alteration of a bit position, and with other codes a variety of diversity-generating operators may be used. By itself, mutation induces a simple random walk through string space. When it is used with selection alone, the two combine to form a parallel, noise-tolerant, hill-climbing algorithm. When used together with selection and crossover, mutation acts both as an insurance policy against losing needed diversity and as a hill climber.

Capabilities. Genetic algorithms can be attractive for a number of reasons: they can solve hard problems quickly and reliably, are easy to interface to existing simulations and models, are extensible, and are easy to hybridize.

One of the primary reasons to use genetic algorithms is that they are efficient and effective over a broad class of problems. Empirical work has long suggested this characteristic, but theory is catching up, and it appears that genetic algorithms can quickly solve problems that have many difficult-to-find optima. There is now a considerable body of genetic algorithm theory from application design to fully rigorous nonlinear difference equations and Markov chains. Application design theory gives useful, yet simple, algebraic relations for the principled selection of population sizes and other algorithm parameters, eliminating unnecessary trial and error. This understanding when properly used to design and size scalable genetic algorithms, sometimes called competent genetic algorithms, results in procedures that appear to solve a large class of difficult problems in times that grow as a subquadratic function of the number of decisions variables in a probably approximately correct sense. *See* STOCHASTIC PROCESS.

Because genetic algorithms require very little problem-specific information, they are remarkably easy to connect to extant application code. Many algorithms require a high degree of integration between solver and objective function. For most users, the difficult part of optimization is getting a good model, and once it is tested and calibrated, the genetic algorithm can be interfaced quite directly without additional difficulty.

Even simple genetic algorithms can be broadly capable, but real problems can pose unanticipated difficulties. When these problems arise, there is often a solution from nature available to solve the problem. For example, many problems in artificial intelligence have search spaces that are highly multimodal and solution sets that may contain multiple global solutions. In these cases, it is desirable to have a population converge to multiple optima simultaneously. By borrowing the biological notions of niche and speciation, genetic algorithms can be designed to converge to a diverse set of highly fit solutions, thereby solving multimodal problems in a single run.

When no solution from nature is available, it is often possible to use problem-specific information to help make a hybrid or knowledge-augmented genetic algorithm. Effective hybrids can be constructed by combining the global perspective of the genetic algorithm with the local efficiency of a problem-specific heuristic. There are also a number of ways that problem-specific information can be built into the operators or the codings. Hybrids are one effective means of enhancing the efficiency of genetic algorithms. Researchers have been exploring effective parallelization, time utilization, and evaluation relaxation as well. Applicable design theory may be used in each of these areas to construct principled approaches to acceptable solutions in a reasonable time.

Applications. Applications have fallen into one of two categories: search and optimization, and machine learning. Genetic algorithms have also inspired other forms of computation and novel scientific approaches.

Search and optimization. Some of the notable commercial applications of genetic algorithms in search and optimization include gas and steam turbine design, expansion of fiber-optic communications networks, and automatic drawings of criminal suspect faces from witness recollection. Prospective applications include financial time-series prediction, protein folding, aerospace composite structural design, and integrated circuit design. Two important trends in the application of genetic algorithms are the use of multiobjective and interactive genetic algorithms. Multiobjective genetic algorithms exploit the existence of a population to search for nondominated (Pareto) solutions along two or more objectives. Interactive genetic algorithms combine human judgment with computerized innovation to permit the solution of problems where the creation of an objective function is difficult or impossible. Interactive genetic algorithms have been used to create pleasing art, music, and advertising, to name a few examples. *See* OPTIMIZATION.

Machine learning. Some of the more interesting, and difficult, applications of the ideas of evolutionary computation come under the heading of genetics-based machine learning (GBML). Perhaps the most ambitious of GBML systems is the learning classifier system, a type of learning expert system that combines a primitive expert system shell, genetic and other learning algorithms, and a means of apportioning credit to alternative rules. *See* EXPERT SYSTEMS.

A more direct approach to solving GBML problems has arisen in genetic programming systems. In genetic programming, modified simple genetic algorithms are used to recombine computer programs, which are then rated in their ability to solve some problem. The more highly fit programs are recombined with others to form new, possibly better programs. A broad array of problems have been tackled by using the methods of genetic programming. Such progress is not surprising since natural genotypes are dynamic structures more akin to computer programs than to static sets of decision-variable values. *See* PROGRAMMING LANGUAGES.

Computational and scientific influences. Genetic algorithms have been instrumental for creating other forms of biologically inspired computation. For example, ant colony optimization, artificial immune systems, and particle swarm optimization use different metaphors as motivation, but their population orientation and detailed mechanics have much in common with, and have drawn heavily from, evolutionary computation.

Genetic algorithms have also influenced a number of important scientific directions. Agent-based computational economics, artificial life, and other forms of complex adaptive systems inquiry draw heavily from the example, theory, and lore of genetic algorithms. These approaches now are being used to create computational laboratories in scientific fields where experimentation is otherwise difficult. *See* ALGORITHM; ARTIFICIAL INTELLIGENCE; INTELLIGENT MACHINE. **David E. Goldberg**

Bibliography. T. Baeck, D. B. Fogel, and M. Michalewicz (eds.), *Handbook of Evolutionary Computation*, 1997; K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*, 2001; D. E. Goldberg, *The Design of Innovation: Lessons from and for Competent Genetic Algorithms*, 2002; D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, 1989; J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, 1992; J. H. Holland, *Hidden Order: How Adaptation Builds Complexity*, 1996; C. L. Karr and L. M. Freeman (eds.), *Industrial Applications of Genetic Algorithms*, 1998; J. Koza et al., *Genetic Programming 3: Darwinian Invention and Problem Solving*, 1999; W. B. Langdon et al., *Foundations of Genetic Programming*, 2001.

Genetic code

The rules by which the base sequences of deoxyribonucleic acid (DNA) are translated into the amino acid sequences of proteins. Each sequence of DNA that codes for a protein is transcribed or copied (**Fig. 1**) into messenger ribonucleic acid (mRNA). Following the rules of the code, discrete elements in the mRNA, known as codons, specify each of the 20 different amino acids that are the constituents of proteins. In a process called translation, the cell decodes the message in mRNA. During translation (**Fig. 2**), another class of RNAs, called transfer RNAs (tRNAs), are coupled to amino acids, bind to the mRNA, and in a step-by-step fashion provide the amino acids that are linked together in the order called for by the mRNA sequence. The specific attachment of each amino acid to the appropriate tRNA, and the precise pairing of tRNAs via their anticodons to the correct codons in the mRNA, form the basis of the genetic code. *See* DEOXYRIBONUCLEIC ACID (DNA); PROTEIN; RIBONUCLEIC ACID (RNA).

Universal genetic code. The genetic information in DNA is found in the sequence or order of four bases that are linked together to form each strand of the two-stranded DNA molecule. The bases of DNA are adenine, guanine, thymine, and cytosine, which are abbreviated A, G, T, and C. Chemically, A and G are purines, and C and T are pyrimidines. The two strands of DNA are wound about each other in a double helix that looks like a twisted ladder (Fig. 1). The two posts of the ladder consist of sugar (deoxyribose) and phosphate chains that link the bases together in a directional manner. In DNA, the two strands are said to be antiparallel, because they run in opposite directions. Each rung of the ladder is formed by two bases, one from each strand, that pair with each other by means of hydrogen bonds. For a good fit, a pyrimidine must pair with a purine; in DNA, A bonds with T, and G bonds with C. *See* PURINE; PYRIMIDINE.

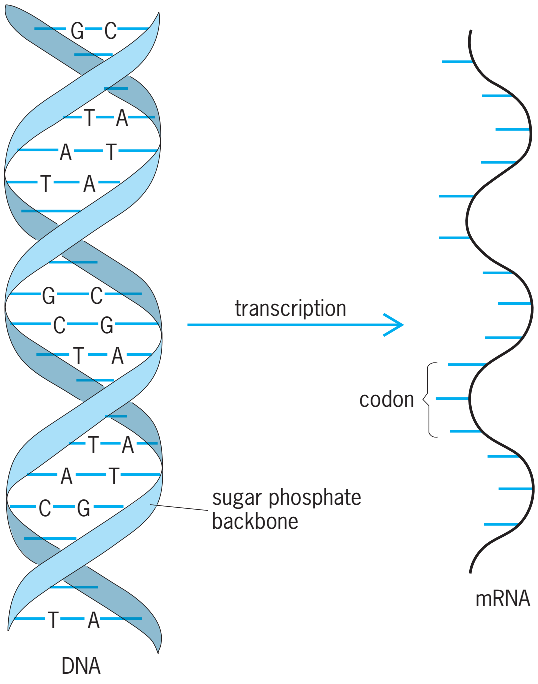


Fig. 1. DNA double helix and mRNA copy made from it by transcription.

Ribonucleic acids such as mRNA or tRNA also comprise four bases, except that in RNA the pyrimidine uracil (U) replaces thymine. During transcription a

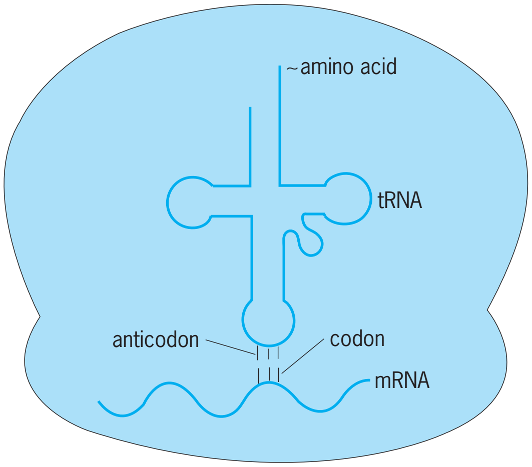


Fig. 2. Translation process in the ribosome, showing the tRNA with attached amino acid interacting with the mRNA via the codon-anticodon interaction.

single-stranded mRNA copy of one strand of the DNA is made. That mRNA copy of the DNA is bound to a large particle known as the ribosome and decoded according to the rules of the genetic code (Figs. 1 and 2). *See* RIBOSOMES.

If two bases at a time are grouped together, then only 4 × 4 or 16 different combinations are possible, a number that is insufficient to code for all 20 amino acids that are found in proteins. However, if the four bases are grouped together in threes, then there are 4 × 4 × 4 or 64 different combinations. Read sequentially without overlapping, those groups of three bases constitute a codon, the unit that codes for a single amino acid.

The 64 codons can be divided into 16 families of four (**Fig. 3**), in which each codon begins with the same two bases. With the number of codons exceeding the number of amino acids, several codons can code for the same amino acid. Thus, the code is degenerate. In eight instances, all four codons in a family specify the same amino acid. In the remaining families, the two codons that end with the pyrimidines U and C often specify one amino acid, whereas the two codons that end with the purines A and G specify another. Furthermore, three of the codons, UAA, UAG, and UGA, do not code for any amino acid but instead signal the end of the protein chain.

Decoding mRNAs with tRNAs. The nucleic acid code of an mRNA is converted into an amino acid sequence with the aid of tRNAs. These RNAs are relatively small nucleic acids that vary from 75 to 93 bases in length. Some of the bases in a tRNA pair with each other, so that a tRNA is folded into a pattern that can be represented in two dimensions as a cloverleaf (Fig. 2). This pattern is further folded in three dimensions to form an L-shaped molecule (**Fig. 4**), to which an amino acid can then be attached. At the other end, three bases in a single-stranded loop are free to pair with a codon in the mRNA. These three bases of a tRNA constitute the anticodon.

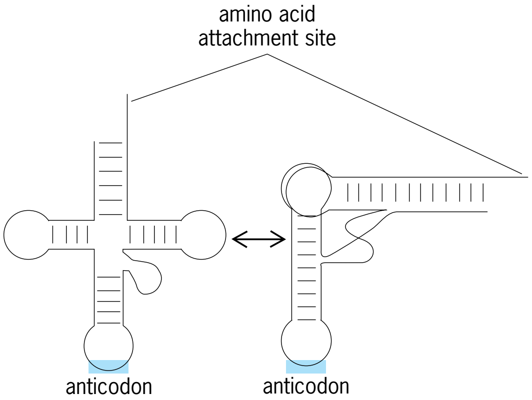
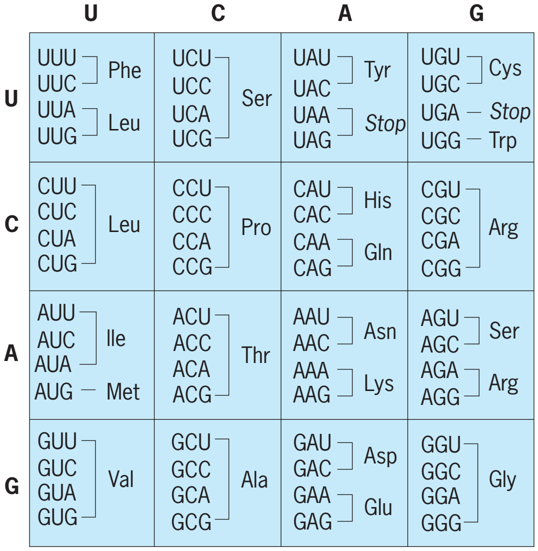


Fig. 4. A transfer RNA in the cloverleaf representation is shown with a two-dimensional projection of how the structure is folded into an L-shaped molecule in three dimensions. The short horizontal and vertical lines represent base pairs that form within the structure. (*Courtesy of Dr. Douglas Buechter*)



	U	C	A	G
U	UUU, UUC Phe UUA, UUG Leu	UCU, UCC, UCA, UCG Ser	UAU, UAC Tyr UAA, UAG *Stop*	UGU, UGC Cys UGA – *Stop* UGG – Trp
C	CUU, CUC, CUA, CUG Leu	CCU, CCC, CCA, CCG Pro	CAU, CAC His CAA, CAG Gln	CGU, CGC, CGA, CGG Arg
A	AUU, AUC, AUA Ile AUG – Met	ACU, ACC, ACA, ACG Thr	AAU, AAC Asn AAA, AAG Lys	AGU, AGC Ser AGA, AGG Arg
G	GUU, GUC, GUA, GUG Val	GCU, GCC, GCA, GCG Ala	GAU, GAC Asp GAA, GAG Glu	GGU, GGC, GGA, GGG Gly

Fig. 3. Universal (standard) genetic code. Each of the 64 codons found in mRNA specifies an amino acid (indicated by the three-letter abbreviation) or the end of the protein chain (stop). The amino acids are phenylalanine (Phe), leucine (Leu), isoleucine (Ile), methionine (Met), valine (Val), serine (Ser), proline (Pro), threonine (Thr), alanine (Ala), tyrosine (Tyr), histidine (His), glutamine (Gln), asparagine (Asn), lysine (Lys), aspartic acid (Asp), glutamic acid (Glu), cysteine (Cys), tryptophan (Trp), arginine (Arg), and glycine (Gly).

Each amino acid has one or more tRNAs, and because of the degeneracy of the code, many of the tRNAs for a specific amino acid have different anticodon sequences. However, the tRNAs for one amino acid are capable of pairing their anticodons only with the codon or codons in the mRNA that specify that amino acid. The tRNAs act as interpreters of the code, providing the correct amino acid in response to each codon by virtue of precise codon–anticodon pairing. The tRNAs pair with the codons and sequentially insert their amino acids in the exact order specified by the sequence of codons in the mRNA. That takes place in the ribosome.

Codon-anticodon interaction. The anticodon of the tRNA pairs with the codon in the mRNA in an antiparallel fashion so that the first base of the codon pairs with the third base of the anticodon, the second base of the codon pairs with the second base of the anticodon, and the third base of the codon pairs with the first base of the anticodon. For the first two bases of the codon, pairing follows the usual rule: A pairs with U, and G with C. However, at the third base of codon, pairing occurs according to a set of rules first suggested by F. Crick and known as the wobble hypothesis:

First base of anticodon	*Third base of codon*
C	G
U	A or G
G	C or U
I	U, C, or A

The purine base inosine (I) is often found instead of A in the first position of the anticodon. There, inosine can pair with any of three bases in the third position of the codon, thus allowing a single tRNA to read three out of the four codons for a single amino acid. However, inosine is not found in tRNAs in which it would allow pairing with a codon that specifies another amino acid. For example, the histidine tRNA does not contain inosine in the first position of the anticodon because that would allow pairing with the CAA codon which codes for glutamine, not histidine.

Variations. The rules of the genetic code are virtually the same for all organisms, but there are some interesting exceptions. In the microorganism *Mycoplasma capricolum*, UGA is not a stop codon; instead it codes for tryptophan. This alteration in the code is also found in the mitochondria of some organisms. In contrast, UGA is the sole termination codon in some ciliated protozoa, where UAA and UAG are translated as glutamine. A special case involves the unusual amino acid selenocysteine. This metal (selenium)–containing amino acid has a unique tRNA that inserts selenocysteine in response to a UGA codon, in a limited number of proteins in bacteria and mammals. Because these organisms also use UGA as a stop codon, it is thought that the base sequence of the mRNA around selenocysteine-encoding UGAs influences reading to favor amino acid insertion instead of termination. Other specific changes in the meaning of codons that normally encode isoleucine, leucine, and arginine have also been observed in the mitochondria of some organisms.

In addition to changes in the meanings of codons, a modified system for reading codons that requires fewer tRNAs is found in mitochondria. Following the wobble rules, a minimum of two tRNAs are normally needed to read each family of four codons. For example, two tRNAs, with I or C in the first position of the anticodon, could read all four members of a family. In contrast, each of the eight four-codon families that codes for a single amino acid in mitochondria is read by a single tRNA with U in the first position of the anticodon. An explanation for this apparent relaxation of the wobble rules is that the presence of a chemical modification of U in the universal code restricts pairing of U to A or G. In mitochondrial tRNAs, U is unmodified and is therefore able to pair with codons ending in U, C, A, or G. In those cases where two codons in a family specify one amino acid and the other two members specify another amino acid or a stop codon (Fig. 3), the U in the first position of the anticodon is modified and thus pairs only with A and G, as in the universal system. *See* MITOCHONDRIA.

Aminoacyl-tRNA synthetases. The specific attachment of amino acids to tRNAs is catalyzed by a class of enzymes known as aminoacyl-tRNA synthetases. Each amino acid has a different aminoacyl-tRNA synthetase, which catalyzes a reaction known as aminoacylation. Each enzyme attaches its particular amino acid to all of the tRNAs that have anticodons corresponding to that amino acid. Each aminoacyl-tRNA synthetase must distinguish between the various tRNAs so that only the proper ones are aminoacylated. Thus, the accurate recognition of tRNAs by these enzymes establishes the rules of the genetic code.

Aminoacyl tRNA synthetases are divided into two classes of 10 synthetases each. The classification of a synthetase is based on which of two distinct three-dimensional structures it has for the active-site region that catalyzes attachment of an amino acid to its cognate tRNA. There is no evidence that the two distinct structures for the active-site region came from a common ancestor, and it is believed that the two classes arose independently. Also, the occurrence of two classes of synthetases appears to be universal; that is, there is no known example of an organism where all synthetases are of the same class. The historical rationale for the two classes of synthetases is a question of great interest. An attractive hypothesis that is gaining support postulates that, in an early environment, synthetases from opposite classes paired themselves on a single tRNA. One member from each class bound to opposite sides of a given tRNA to form a two-synthetase–one-tRNA complex. From this perspective, the distinct architectures of the two classes of synthetases are seen as needed in order to fit two synthetases to distinct sides of the tRNA.

Because synthetases are responsible for the genetic code and are present in all living organisms, it is likely that they are among the oldest enzymes on Earth. The active-site region of a synthetase is thought to be the primordial, or earliest, form of the synthetase. In addition to the active-site-containing region that is conserved among all members of the same class, each synthetase has a region that is referred to as a nonconserved domain. This domain is more or less unique to the synthetase, even within a particular class.

Recognition of tRNAs. Aminoacyl tRNA synthetases engage in a process called recognition of tRNAs, analogous to the recognition of different buildings or trees by the human eye. Each tRNA has one or more distinguishing characteristics that are determined by the nucleotide sequence and associated three-dimensional structure. The distinguishing features are recognized by the appropriate synthetase that catalyzes the attachment of its cognate amino acid.

One feature that distinguishes a particular tRNA is the anticodon trinucleotide. Were the anticodon trinucleotide the only feature that synthetases use to distinguish among tRNAs, then the interaction of a tRNA with a synthetase would establish a direct connection between a particular amino acid and its complementary trinucleotide of the genetic code. However, there are synthetases that do not interact with the anticodon and yet still correctly identify their tRNA. In these cases, the relationship between a particular trinucleotide and a given amino acid is indirect, and suggests the existence of another code, namely, an operational RNA code for amino acids that was the precursor to the genetic code.

Second genetic code. In the case of the tRNA specific for alanine with an anticodon GGC, a single G-U base pair, at a specific position far removed from the anticodon, is needed for aminoacylation with alanine. This base pair is located in a portion of the tRNA molecule relatively close to the site of amino acid

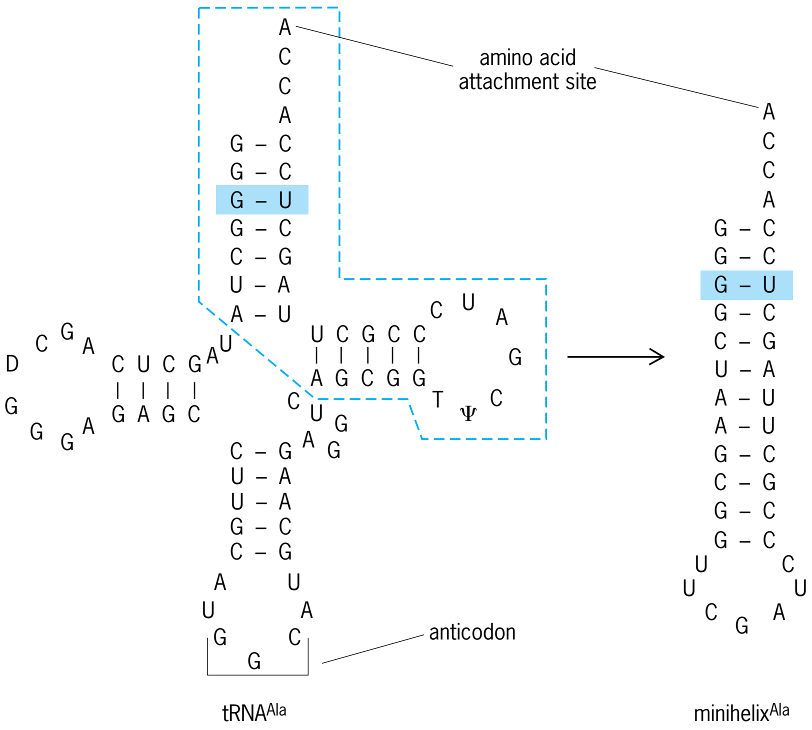


Fig. 5. Cloverleaf representation of the nucleotide sequence, comprising four monomer units (A = adenine, G = guanine, C = cytosine, and U = uracil), of an alanine tRNA with a GGC anticodon and of a synthetic minihelix based on a limited portion of the whole tRNA (enclosed by a broken line). D, T, and Ψ are modified uracil bases in the tRNA, and T and Ψ have been replaced by U in the minihelix. The critical G-U base pair recognized by the alanine tRNA synthetase is highlighted in both structures. (*Courtesy of Dr. Douglas Buechter*)

attachment (**Fig. 5**). The part of the tRNA structure that is near the amino acid attachment site is known as the acceptor stem, and consists of a short helix joined to a single-stranded piece that terminates at the site of amino acid attachment. A truncated version of alanine tRNA that contains the critical base pair together with the adjacent structure near the amino acid attachment site can be synthesized as a minihelix. This tRNA fragment is recognized by alanine tRNA synthetase.

Even smaller RNA fragments that recapitulate the acceptor ends of several tRNAs are aminoacylated with the correct amino acid (**Fig. 6**). Because aminoacylation is specific and depends on the order of bases in the short fragments, the relationship between the structure-sequence of RNA fragments and each amino acid can be viewed as an operational RNA code, referred to also as the second genetic code.

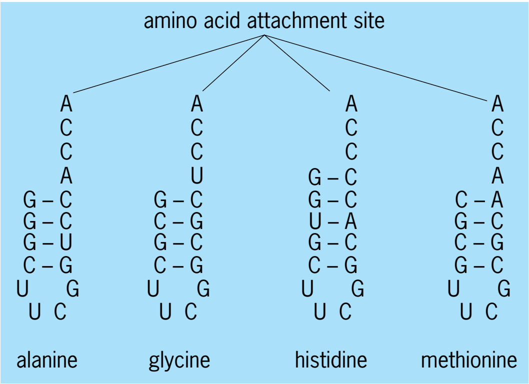


Fig. 6. The smallest RNA oligonucleotides that have been shown to be specifically aminoacylated. The amino acid which can be attached to each RNA oligonucleotide is indicated. (*Courtesy of Dr. Barry Henderson*)

Several synthetases recognize the anticodon as one of the elements needed for identification of the tRNA, but even in these cases nucleotides in the acceptor stem are still important for precise and efficient recognition of the whole tRNA and small RNA fragments. The class-defining and active-site-containing domain of aminoacyl tRNA synthetases is largely responsible for interactions with the acceptor end of a tRNA and, therefore, for interpretation of the second genetic code. After establishment of the second genetic code in a primitive world, RNA fragments with signals for attachment of specific amino acids may have become incorporated into RNA structures which eventually became transfer RNAs. That step, in turn, gave birth to the modern genetic code. In addition, the synthetases developed editing functions that are tRNA-dependent. The development of these editing functions was one of the events that linked the evolution of tRNAs with that of the synthetases. These editing functions enable the synthetases to achieve high accuracy of aminoacylation. During evolution, the earliest tRNA synthetases apparently acquired additional, unique domains which allowed for editing. *See* GENE; GENE ACTION; GENETICS.

P. Schimmel; K. Ewalt

Bibliography. J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed., 2002; S. Cusack, RNA-protein complexes, *Curr. Opin. Struc. Biol.*, 9:66–73, 1999; A. J. Griffiths, *An Introduction to Genetic Analysis*, 6th ed., 1999; P. Schimmel and L. Ribas de Pouplana, Formation of two classes of tRNA synthetases in relation to editing functions and genetic code, *Cold Spring Harbor Symp. Quantitative Biol.*, 66:161–166, 2001; P. Schimmel and L. Ribas de Pouplana, Transfer RNA: From minihelix to genetic code, *Cell*, 81:983–986, 1995.

Genetic engineering

The artificial recombination of nucleic acid molecules in the test tube, their insertion into a virus, bacterial plasmid, or other vector system, and the subsequent incorporation of the chimeric molecules into a host organism in which they are capable of continued propagation. The construction of such molecules has also been termed gene manipulation because it usually involves the production of novel genetic combinations by biochemical means. *See* NUCLEIC ACID.

Genetic engineering provides the ability to propagate and grow in bulk a line of genetically identical organisms, all containing the same artificially recombinant molecule. Any genetic segment as well as the gene product encoded by it can therefore potentially be amplified. For these reasons the process has also been termed molecular cloning or gene cloning. *See* GENE.

Basic Techniques

The central techniques of such gene manipulation involve (1) the isolation of a specific deoxyribonucleic acid (DNA) molecule or molecules to be replicated (the passenger DNA); (2) the joining of this DNA with a DNA vector (also known as a vehicle or a replicon) capable of autonomous replication in a living cell after foreign DNA has been inserted into it; and (3) the transfer, via transformation or transfection, of the recombinant molecule into a suitable host.

Isolation of passenger DNA. Passenger DNA may be isolated in a number of ways; the most common of these involves DNA restriction.

DNA restriction. DNA purified from an organism can be prepared for cloning only after it has been cut into smaller molecules. Restriction endonucleases make possible the cleavage of high-molecular-weight DNA. Although three different classes of these enzymes have been described, only type II restriction endonucleases have been used extensively in the manipulation of DNA. Type II restriction endonucleases are DNAases that recognize specific short nucleotide sequences (usually 4 to 6 base pairs in length), and then cleave both strands of the DNA duplex, generating discrete DNA fragments of defined length and sequence. *See* RESTRICTION ENZYME.

A number of restriction enzymes make staggered cuts in the two DNA strands, generating single-stranded termini. Restriction enzyme Eco R1 cleaves the sequence GAATTC (where G = guanine; A = adenine; T = thymine; C = cytosine) between the G and A residues, generating protruding single-stranded ends having 5′ termini.

The various fragments generated when a specific DNA is cut by a restriction enzyme can be easily resolved as bands of distinct molecular weights by agarose gel electrophoresis. Specific sequences of these bands can be identified by a technique known as Southern blotting. In this technique, DNA restriction fragments resolved on a gel are denatured and blotted onto a nitrocellulose filter. The filter is incubated together with a radioactively labeled DNA or RNA probe specific for the gene under study. The labeled probe hybridizes to its complement in the restricted DNA, and the regions of hybridization are detected autoradiographically. Fragments of interest can then be eluted out of these gels and used for cloning. Purification of particular DNA segments prior to cloning reduces the number of recombinants that must later be screened.

Mechanical shearing of DNA. Another method that has been used to generate small DNA fragments is mechanical shearing. Intense sonification of high-molecular-weight DNA with ultrasound, or high-speed stirring in a blender can both be used to produce DNA fragments of a certain size range. Shearing results in random breakage of DNA, producing termini consisting of short, single-stranded regions.

Other sources. Another source of DNA to be replicated is DNA complementary to poly(A) RNA, or cDNA, which is synthesized in the test tube. The cDNA synthesis involves the use of the enzyme, reverse transcriptase, which uses poly(A) RNA as a template for the synthesis of a complementary DNA strand. The RNA strand in the resulting RNA–DNA hybrid is removed by alkaline hydrolysis, and the remaining single-stranded DNA is converted into the double-stranded form in a second reaction involving DNA polymerase I. *See* ENZYME.

In addition, short oligonucleotides, synthesized chemically, can be used for cloning purposes. *See* OLIGONUCLEOTIDE.

Joining DNA molecules. Once the proper DNA fragments have been obtained, they must be joined. This is accomplished by the following methods.

Cohesive end ligation. When cleavage with a restriction endonuclease creates cohesive ends, these can be annealed with a similarly cleaved DNA from another source, including a vector molecule. When such molecules associate, the joint has nicks a few base pairs apart in opposite strands. The enzyme DNA ligase can then repair these nicks to form an intact, duplex recombinant molecule, which can be used for transformation and the subsequent selection of cells containing the recombinant molecule (**Fig. 1**).

Cohesive ends can also be created by the addition of synthetic DNA linkers to blunt-ended DNA molecules. These linkers are short DNA duplexes containing the recognition site for a restriction enzyme which produces cohesive termini. Linkers are ligated to blunt-ended passenger DNA molecules by DNA ligase encoded by the phage T4. After digesting the product with the restriction enzyme that cleaves the linkers, the products can be ligated to any vehicle DNA via the complementary termini and then cloned.

Homopolymer tailing. Another method for joining DNA molecules involves the addition of homopolymer extensions to different DNA populations followed by an annealing of complementary homopolymer sequences. For example, short nucleotide sequences of pure adenine can be added to the 3′ ends of one population of DNA molecules and short thymine blocks to the 3′ ends of another population. The two types of molecules can then anneal to form mixed dimeric circles that can be used directly for transformation. Single-stranded gaps that may remain in the two strands at each connection will be repaired in the transformed cell.

Blunt-end ligation. T4 DNA ligase carries out the intermolecular joining of DNA substrates at completely base-paired ends; such blunt ends can be produced by cleavage with a restriction enzyme or by mechanical shearing followed by enzyme treatment.

Transformation. The desired DNA sequence, once attached to a DNA vector, must be transferred to a suitable host. Transformation is defined as the introduction of foreign DNA into a recipient cell. Transformation of a cell with DNA from a virus is usually referred to as transfection.

Transformation in any organism involves (1) a method that allows the introduction of DNA into

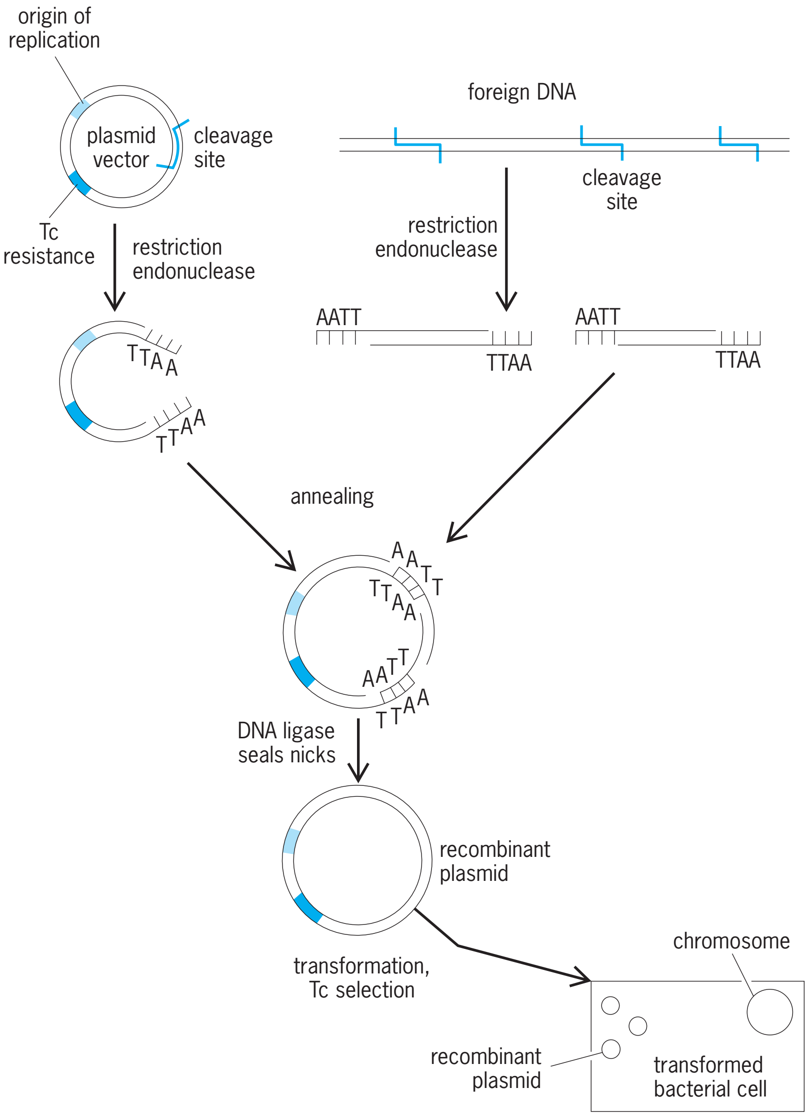


Fig. 1. Use of DNA ligase to create a covalent DNA recombinant joined through association of termini generated by Eco R1. (*After R. W. Old and S. B. Primrose, Principles of Gene Manipulation, 3d ed., Blackwell Scientific Publications, 1985*)

the cell and (2) the stable integration of DNA into a chromosome, or maintenance of the DNA as a self-replicating entity. *See* TRANSFORMATION (BACTERIA).

Prokaryotes and lower eukaryotes. *Escherichia coli* is usually the host of choice for cloning experiments, and transformation of *E. coli* is an essential step in these experiments. *Escherichia coli* treated with calcium chloride are able to take up DNA from bacteriophage lambda as well as plasmid DNA. Calcium chloride is thought to effect some structural alterations in the bacterial cell wall.

In addition to *E. coli*, other prokaryotes have also been used as hosts for cloning experiments, among them the nonenteric bacteria, *Bacillus* species, and actinomycetes. An efficient method for transformation in *Bacillus* involves poly(ethylene glycol)–induced DNA uptake in bacterial protoplasts and subsequent regeneration of the bacterial cell wall. Actinomycetes can be similarly transformed. Transformation can also be achieved by first entrapping the DNA with liposomes followed by their fusion with the host cell membrane. Similar transformation methods have been developed for lower eukaryotes such as the yeast *Saccharomyces cerevisiae* and the filamentous fungus *Neurospora crassa*. *See* LIPOSOMES.

Animal cells. Several methods are available for the transfer of DNA into cells of higher eukaryotes. Specific genes or entire viral genomes can be introduced into cultured mammalian cells in the form of a coprecipitate with calcium phosphate. DNA complexed with calcium phosphate is readily taken up and expressed by mammalian cells. DNA complexed with diethylamino-ethyl-dextran (DEAE-dextran), or DNA trapped in liposomes or erythrocyte ghosts may also be used in mammalian transformation. Alternatively, bacterial protoplasts containing plasmids can be fused to intact animal cells with the aid of chemical agents such as poly(ethylene glycol) [PEG]. Finally, DNA can be directly introduced into cells by microinjection. The efficiency of transfer by each of these methods is quite variable.

Higher plants. Introduction of DNA sequences by insertion into the transforming (T)-DNA region of the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* is a method of introducing DNA into plant cells and ensuring its integration. The Ti plasmid has been used to introduce selectable marker genes, including bacterial genes coding for antibiotic resistance, into plant cells. The control sequences used to express these chimeric genes are derived from plant genes. Many foreign genes have been expressed in the transformant plant cells and have conferred selectable properties into plant cells.

Although the Ti plasmid can be used as a gene transfer vector, its use is limited in that *A. tumefaciens*, the bacterium that transmits Ti plasmids, normally infects primarily dicotyledonous plants. Monocotyledonous plants, including economically important cereals, cannot be infected by the Ti plasmid.

Because of the limitations of the host range of *A. tumefaciens*, alternative transformation systems are being developed for gene transfer in plants. They include the use of liposomes, as well as induction of DNA uptake in plant protoplasts. With those methods, however, transformation frequencies are low. Plant DNA viruses, such as the cauliflower mosaic virus, might also be used as vectors for DNA transfer.

Foreign DNA has been introduced into plant cells by a technique called electroporation. This technique involves the use of electric pulses to make plant plasma membranes permeable to plasmid DNA molecules. Plasmid DNA taken up in this way has been shown to be stably inherited and expressed.

Cloning Vectors

There is a large variety of potential vectors for cloned genes. The vectors differ in different classes of organisms.

Prokaryotes and lower eukaryotes. Three types of vectors have been used in these organisms: plasmids, bacteriophages, and cosmids.

Plasmids. Plasmids are extrachromosomal DNA sequences that are stably inherited. *Escherichia coli* and its plasmids constitute the most versatile type of host–vector system known for DNA cloning.

Several natural plasmids, such as ColE1, have been used as cloning vehicles in *E. coli*. In addition, a variety of derivatives of natural plasmids have been constructed by combining DNA segments and desirable qualities of older cloning vehicles. These qualities, which facilitate manipulation of cloned DNA and limit the ability of recombinants to survive in the environment beyond the laboratory, include (1) reduced molecular weight of the plasmid; (2) construction of only one site for a given restriction enzyme located in a region not essential for plasmid replication; (3) introduction of one or more phenotypic markers into the plasmid to permit selection of cells containing the plasmid and to aid in the identification of plasmids that have incorporated the foreign DNA. It is also desirable to use plasmids that can be obtained in high yields.

The most versatile and widely used of these plasmids is pBR322. This plasmid has been completely sequenced, providing a detailed restriction enzyme cleavage map and DNA sequence information (**Fig. 2**). The pBR322 plasmid carries the genes coding for resistance to the antibiotics ampicillin and tetracyline.

Transformation in yeast has been demonstrated using a number of plasmids, including vectors derived from the naturally occurring 2μ plasmid of yeast. These vectors contain DNA fragments from the 2μ yeast plasmid, yeast nuclear DNA, and the *E. coli* vector, pMB9. Other yeast vectors are recombinants consisting of *E. coli* vectors and yeast DNA fragments.

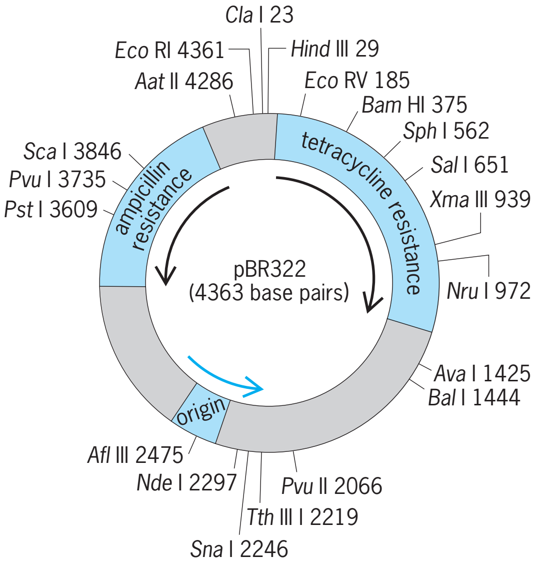


Fig. 2. Structure of pBR322 showing the unique cleavage sites. The numbers relate to the coordinate of the base which corresponds to the 5′ nucleotide of each recognition sequence. The black arrows show the directions of transcription of the Ap^R and Tc^R genes. The colored arrow shows the direction of DNA replication. (*After R. W. Old and S. B. Primrose, Principles of Gene Manipulation, 3d ed., Blackwell Scientific Publications, 1985*)

Bacteriophages. Bacteriophage lambda is a virus of *E. coli*. Several lambda-derived vectors have been developed for cloning in *E. coli*, and for the isolation of particular genes from eukaryotic genomes. These lambda derivatives have several advantages over plasmids: (1) Thousands of recombinant phage plaques can easily be screened for a particular DNA sequence on a single petri dish by molecular hybridization. (2) Packaging of recombinant DNA in laboratory cultures provides a very efficient means of DNA uptake by the bacteria. (3) Thousands of independently packaged recombinant phages can be easily replicated and stored in a single solution as a "library" of genomic sequences. *See* BACTERIOPHAGE.

Cosmids. Plasmids have also been constructed that contain the phage *cos* DNA site, required for packaging into the phage particles, and ColE1 DNA segments, required for plasmid replication. These plasmids have been termed cosmids. After packaging laboratory cultures, the vector is used to infect a host. The recombinant cosmid DNA is injected and circularizes like phage DNA but replicates as a plasmid. Transformed cells are selected on the basis of a vector drug resistance marker. Cosmids provide an efficient means of cloning large pieces of foreign DNA and have therefore been used for constructing libraries of eukaryotic genomic fragments.

Animal cells. In contrast to the wide variety of plasmid and phage vectors available for cloning in prokaryotic cells, relatively few vectors are available for introducing foreign genes into animal cells. In the latter case, the most commonly used vectors are derived from simian virus 40 (SV40). *See* ADENO-SV40 HYBRID VIRUS.

The SV40 genome consists of a small covalently closed circular DNA molecule for which the entire nucleotide sequence is known. Normal SV40 cannot be used as a vector, since there is a physical limit to the amount of DNA that can be packaged into the virus capsid, and the addition of foreign DNA would generate a DNA molecule too large to be packaged into a viral particle. However, SV40 mutants lacking portions of the genome can be propagated in mixed infections in which a "helper" virus supplies the missing function.

Propagating recombinants in eukaryotic cells using SV40 vectors imposes some limitations, such as the severe limitation on the size of gene segments that can be cloned in the small viral genome and the fact that only cells permissive for SV40 replication can be used.

Bovine papilloma virus is another DNA virus that has been used as a vector in eukaryotic animal cells. The rat insulin gene inserted into the viral DNA has been used successfully to transform mouse cells.

Another system for the stable transfer of genes into mammalian cells involves cotransformation of cells with two physically separate sets of genes. For example, the introduction of the thymidine kinase (tk) gene from Herpes simplex virus to mouse tk^- cells results in stable and efficient transformation of cells that express the viral gene.

In *Drosophilia*, transposable DNA elements, that is, DNA segments that are capable of moving from one position to another within the genome of a cell, have provided the basis for development of an efficient and controlled system of gene transfer. DNA segments of interest can be transposed into germ-line chromosomes along with a transposable element to which the segment has been previously ligated. *See* TRANSPOSONS.

Plant cells. Two systems for the delivery and integration of foreign genes into the plant genome are the Ti plasmid of the soil bacterium *Agrobacterium* and the DNA plant virion cauliflower mosaic virus.

The Ti plasmid is a natural gene transfer vector carried by *A. tumefaciens*, a pathogenic bacterium that causes crown gall tumor formation in dicotyledonous plants. A T-DNA segment present in the Ti plasmid becomes stably integrated into the plant cell genome during infection. This property of the Ti plasmid has been exploited to show that DNA segments inserted in the T-DNA region can be cotransferred to plant DNA. *See* CROWN GALL.

Caulimoviruses belong to a group of plant DNA viruses that contain double-stranded DNA. Of these, cauliflower mosaic virus is the best studied for its cloning vehicle potential. The genomes of over 30 strains of cauliflower mosaic virus have been mapped by using restriction enzymes. Purified preparations of the cauliflower mosaic virus DNA can infect plant cells with relatively high efficiency. Cauliflower mosaic virus DNA is also transformable after gene manipulation, which indicates that it might be possible to use cauliflower mosaic virus as a plant vector.

Cloned Gene Expression

It is sometimes the aim of the genetic engineer to promote the expression of a cloned gene not only for the analysis of gene structure and function but also for amplification of the synthesis of a desirable gene product. DNA cloning methods have enabled the genetic modification of bacteria and unicellular eukaryotes, rendering them capable of producing virtually any gene product of animal or plant cells.

The aim of gene transfer experiments is to examine the mechanisms of gene regulation and expression in the context of normal cell differentiation. DNA sequences cloned by recombinant DNA technology have been transferred into fertilized mouse, *Drosophila*, and *Xenopus* oocytes by microinjection techniques. In the mouse and *Drosophila* this system has allowed study of the expression and regulation of transferred gene sequences in specific tissues during normal cellular differentiation, and their transmission from one generation to the next. The *Xenopus* oocyte has been used mainly for analysis of transcription and translation of injected sequences.

In the mouse, foreign DNA is injected into a pronucleus of a fertilized egg and subsequently becomes incorporated into the chromosomes of the diploid zygote. The injected eggs are then transferred to a foster-mother mouse, where normal embryonic development occurs. Some of the progeny mice come to contain the foreign DNA in their cells, and the expression of the introduced DNA can be examined during development of these transgenic mice.

A variety of recombinant DNAs have been injected into the mouse oocyte, including plasmids containing SV40 sequences, the Herpes simplex tk gene, a human alpha interferon cDNA, and beta-globin genes from a variety of animals. The injected DNA has been shown to be integrated into the genomes of some of the transformants, and in many cases transmission of DNA from the original transgenic mouse to its progeny has also been demonstrated.

Applications. Genetic engineering technology has made possible a wide variety of practical applications, several of which are discussed below.

Isolation of specific genes. Recombinant DNA technology has permitted the isolation and detailed structural analysis of a large number of prokaryotic and eukaryotic genes. This contribution is especially significant in the eukaryotes because of their large genomes. The methods outlined above provide a means of fractionating and isolating individual genes, since each clone contains a single sequence or a few DNA sequences from a very large genome. Isolation of a particular sequence of interest has been facilitated by the ability to generate a large number of clones and to screen them with the appropriate "probe" (radioactively labeled RNA or DNA) molecules.

Analysis of gene structure and function. Genetic engineering techniques provide pure DNAs in amounts sufficient for mapping, sequencing, and direct structural analyses. Furthermore, gene structure–function relationships can be studied by reintroducing the cloned gene into a eukaryotic nucleus and assaying for transcriptional and translational activities. The DNA sequences can be altered by mutagenesis, before their reintroduction in order to define precise functional regions.

Such analyses have provided insights into the complexity of the eukaryotic genome. For example, many eukaryotic genes contain noncoding sequences (introns) interspersed in the coding sequences (exons) of genes. The primary RNA transcripts are spliced to remove the noncoding sequences as the mRNA matures. Exon–intron boundaries from diverse sources have been sequenced to reveal a common sequence pattern at the opposite ends of different introns. These sequences may be part of a mechanism that brings opposite ends of introns together prior to splicing. *See* EXON; INTRON.

Cloning and sequencing of increasing numbers of eukaryotic genes have brought to light similarities in specific sequences near the 5′ and 3′ ends of different genes. These sequences are suspected to be involved in the regulation of transcription or translation.

Molecular probes for a number of eukaryotic genes have revealed the existence of clustered gene families. Some families, such as the histones, have genes with nearly identical sequences which are arranged tandemly along the chromosomal DNA and which function simultaneously to synthesize their gene products. Members of other gene families, however, though clustered on the chromosome, are not identical; such genes encode closely related proteins. Embryonic and adult globin polypeptides are examples of such a gene family. Within the globin gene cluster are also found "globinlike" sequences that are nonfunctional because of deletions in essential regulatory regions. These segments of DNA may represent vestiges of globin genes that were once functional but, during evolution, have lost their physiological role. *See* HEMOGLOBIN; NUCLEOPROTEIN.

Polypeptide production. Genetic engineering methodology has provided means for the production of polypeptides and proteins. Although this expression of heterologous genes is a function of a variety of complex factors, maximizing the expression of cloned sequences has been under intense and rapid development. It is now possible to produce a wide variety of foreign proteins in *E. coli*. These range from enzymes useful in molecular biology to a vast range of polypeptides with potential human therapeutic applications, such as insulin, interferon, growth hormone, immunoglobins, and enzymes involved in the dynamics of blood coagulation.

Biotechnology on a large scale has evolved as a result of genetic engineering in the laboratory. Human insulin manufactured through the use of genetically engineered *E. coli* has been used safely in diabetes therapy. Other polypeptides such as human growth hormone and interferon have been in production. Research has been under way to manufacture vaccines and many other therapeutic and diagnostic agents by using the new technology. *See* BIOTECHNOLOGY.

Medical applications. The screening of appropriately restricted human DNA with suitable molecular probes, that is, true genes or simply nonrepeated sequences of human DNA cloned in a vector, has disclosed differences between individuals with respect to the size of the restricted DNA fragment containing DNA sequences homologous to the probe. These types of variants within a population, called restriction-fragment-length polymorphisms, can be used to map the human genome. *See* GENETIC MAPPING.

Restriction-fragment-length polymorphisms are inherited by mendelian segregation and are distributed in populations as classical examples of common genetic polymorphisms. If such a DNA variant is located close to a defective gene (which cannot be tested for directly), the DNA variant can be used as a marker to detect the presence of the disease-causing gene. It is estimated that detection of 150 to 300 different DNA markers of this type distributed randomly throughout the human genome would allow detection of any disease-producing gene. Such information would be important in determining individual susceptibility to diseases and in taking measures to prevent them. *See* POLYMORPHISM (GENETICS).

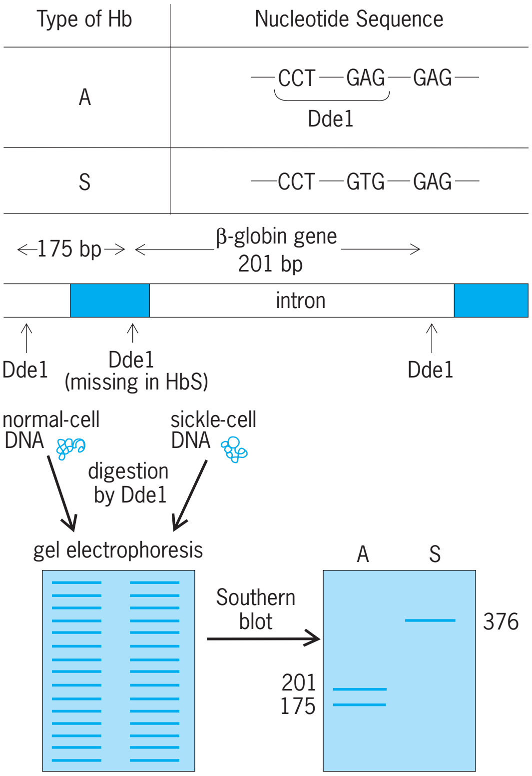


Fig. 3. Detection of the sickle cell globin gene by Southern blotting. The base change (A to T) that causes sickle cell anemia destroys a Dde1 site that is present in the normal β-globin gene. This difference can be detected by digesting sickle cell and normal DNA with Dde1 and performing Southern blot hybridization. Normal DNA will produce two Dde1 fragments of 201 and 175 base pairs, whereas sickle cell DNA will produce a single fragment of 376 base pairs. (*After J. D. Watson et al., Recombinant DNA: A Short Course, W. H. Freeman, 1983*)

In cases where the specific mutation in DNA has been identified, such as sickle cell anemia, certain restriction enzymes that recognize the abnormal DNA sequence at the mutant site can be used to demonstrate the mutation without study of the affected and unaffected family members (**Fig. 3**). By using restriction enzymes in molecular hybridization experiments, changes in DNA sequences that occur in other hemoglobinopathies (genetically determined defects of function or synthesis of hemoglobins) may also be recognized directly by using the appropriate probes. Such changes include deletions, mutations of restriction sites, and deletions of restriction sites. This kind of information provides means of antenatal diagnosis of human diseases by analysis of fetal blood samples.

A number of human diseases are caused by the abnormal function of a major single gene. Various thalassaemias—conditions characterized by reduced or absent synthesis of alpha- or beta-globin chains (the protein components of hemoglobin)—are examples of such diseases. If a normal copy of the defective gene could be introduced and was expressed in the appropriate cells of an individual suffering from such a genetic disorder, the normal gene product would be able to alleviate the problems associated with the presence of the defective gene. Since hemoglobin is produced by certain bone marrow cells that can easily be retrieved, thalassaemia might be treated by the addition of normal isolated genes (in the form of DNA) to the individual's abnormal erythropoietic marrow cells. After the normal genes have been taken up, the manipulated marrow cells could be returned to the person for proliferation and production of normal hemoglobin. *See* HEMATOLOGIC DISORDERS; HEMOGLOBIN.

Genetic modification of plants. Experiments showing the successful transfer and expression of foreign DNA in plant cells using the Ti plasmid, as well as the demonstration that whole plants can be regenerated from cells containing mutated regions of T-DNA, indicate that the Ti plasmid system may be an important tool in the genetic engineering of plants. Such a system will help in the identification and characterization of plant genes as well as provide basic knowledge about gene organization and regulation in higher plants. Once genes useful for crop improvement have been identified, cloned, and stably inserted into the plant genome, it may be possible to engineer plants to be resistant to environmental stress, to pests, and to pathogens. *See* BREEDING (PLANT); GENE; GENE ACTION; SOMATIC CELL GENETICS.

Pamela K. Mulligan

Bibliography. H. Chen, *Medical Genetics Handbook*, 1988; J. Edwards-Moulds and L. C. Lasky (eds.), *Clinical Applications of Genetics Engineering*, 1987; J. Evans and A. Hollaender (eds.), *Genetic Engineering of Animals: An Agricultural Perspective*, 1986; B. R. Glick and J. J. Pasternak, *Molecular Biotechnology: Principles and Applications of Recombinant DNA*, 2d ed., 1998; C. P. Hollenberg and H. Sahm (eds.), *Microbial Genetical Engineering and Enzyme Technology (Biotech 1)*, 1987; D. S. Nicholl, *An Introduction to Genetic Engineering*, 1994; R. Old and S. Primrose, *Principles of Gene Manipulation: An Introduction to Genetic Engineering*, 5th ed., 1994.

Genetic mapping

The resolution of the elements of the genetic material of an organism into a linear array. In order of decreasing size, these elements are chromosomes, genes, and the nucleotide sequence of DNA. A complete description of the genetic material of an organism would consist of its nucleic acid sequence. *See* CHROMOSOME; DEOXYRIBONUCLEIC ACID (DNA).

The construction of a genetic map proceeds through a series of successive approximations. First the chromosomal constitution (karyotype) of the organism is established. Genetically linked loci on each chromosome are grouped and ordered within linkage groups. This process is continued until each chromosome is covered by a set of genetically linked loci to form a complete linkage map. Such maps are used to establish the position of other loci or genes on the chromosome. DNA segments encoding genes on the map are isolated and further analyzed by using the techniques of molecular biology. Finally, the organization and primary DNA structure of genes are determined by DNA sequencing. *See* MOLECULAR BIOLOGY.

The first efforts to sequence DNA, pioneered by Walter Gilbert and Fred Sanger in the 1970s, decoded stretches of DNA measuring a few hundred bases long. The sequencing of the first viral genome (measuring 5000 bases) provided unique insights into the structure and function of genes and its genome organization. The potential to generate vast amounts of information about an organism from its genetic code inspired efforts toword the automation of DNA sequencing and the creation of the Human Genome Project.

The Human Genome Project was conceived in 1985 and implemented in 1990. Its principal goals were to achieve the complete mapping and sequencing of the human genome by 2005. The research program was divided into three principal parts: construction of dense genetic maps, construction of a physical map, and sequencing of the whole genome using the information gained from the first two goals. In 2001 the first draft of the human genome sequence became avaliable to the scientific community, and its completion was announced in April 2003. Since then, the genomic sequences of more than 15 model organisms have also been obtained. *See* HUMAN GENOME PROJECT.

The genetic mapping process has benefited enormously from the resources created through the Human Genome Project. The identification of disease genes has been aided by the availability of dense genetic maps, physical maps, and the genome sequence. The ultimate goal of genetic mapping is to identify disease genes. The isolation and analysis of such genes will provide better tools for early detection of disease susceptibility and eventually could lead to the development of preventive and personalized therapeutic strategies.

Karyotyping and chromosome assignment. The karotype describes the number of complete sets of chromosomes (ploidy) within a cell and the number of chromosomes within each set. For example, human cells are diploid, containing 22 pairs of autosomes (numbered 1 to 22) and one pair of sex chromosomes (XX in females, XY in males). The karyotype of an organism is determined by observation of stained chromosome preparations under a light microscope. Homologous pairs of chromosomes are identified on the basis of size, centromeric position, and the banding pattern produced by such dyes as the Giemsa stain.

Traditionally genes or genetic loci were assigned to a particular human chromosome with the aid of somatic cell hybrids. Fusion of human cells with those of another species, typically a rodent, initially results in hybrid cells containing a complete set of chromosomes from both species. In subsequent cell divisions, human chromosomes are preferentially and randomly lost from such human–rodent hybrids. This results in a series of hybrid cells, each of which contains a subset of the original human chromosome complement. By analyzing a number of such hybrid cells, it is possible to correlate the presence or absence of a particular chromosome with the presence or absence of a particular gene. *See* SOMATIC CELL GENETICS.

The presence of a particular gene within a hybrid cell can be determined in any of a number of ways. In some cases, the human gene may be able to complement a mutation in the recipient cell, or the protein products of equivalent donor and recipient genes can often be distinguished on the basis of size, charge, or enzymatic activity. If the gene of interest encodes proteins present on the cell surface, antibodies raised against these proteins can be used to identify and separate hybrid cells that express such genes from those that do not. This process has been simplified by the use of antibodies with fluorescent labels and a fluorescence-activated cell sorter. *See* ANTIBODY; IMMUNOFLUORESCENCE.

The presence of any DNA sequence can also be detected in hybrid cells by nucleic acid hybridization. Under appropriate conditions, sequence-specific annealing of two complementary single-stranded DNA molecules occurs. A sequence is tested for by the isolation and immobilization of genomic DNA from the hybrid cell on a nylon membrane. Incubation of the membrane with a radioactively labeled single-stranded DNA fragment (probe) in aqueous solution allows the formation of double-stranded molecules between the probe and complementary hybrid DNA sequences. Such duplexes can be detected by autoradiography with x-ray film. Nucleic acid hybridization permits the mapping of any DNA sequence in hybrid cells, including polymorphic DNA markers, independent of the expression of a human gene in hybrid cells. *See* AUTORADIOGRAPHY; NUCLEIC ACID.

If the sequence of the gene is known, a polymerase chain reaction (PCR) assay is used to type the cell hybrid for the relevant sequence. Today, this can also be applied to mapped clones used for sequencing the human genome. In addition, the sequence of the gene in question can be directly compared to and mapped to the human genome. Short sequences can be identified rapidly using the computer program e-PCR.

Longer sequences can be found using the Basic Local Alignment Search Tool (BLAST) or other rapid sequence-similarity searching algorithms.

Isolated DNA sequences can also be mapped to whole chromosomes by in-situ hybridization. At first, in-situ hybridization experiments used radioactive-labeled DNA probes and photographic emulsion for detection. Today, the preferred method is fluorescence in-situ hybridization (FISH), which uses fluorescent detection. Compared with radioactive detection methods, fluorescence detection is safer and faster and can map more than one DNA sequence simultaneously and at a higher resolution.

Somatic cell genetics contributed significantly to our ability to analyze and understand the human genome in the pregenomic era. The mapping methodologies used, such as in-situ hybridization, were applied to the mapping of microbial cloning systems (such as yeast artificial chromosomes, bacterial artificial chromosomes, cosmids, and plasmids), which were used as the bases of modern physical maps.

Linkage mapping. The order and distance between genetic loci on the same chromosome can be determined by the frequency of recombination events between them. Recombination between homologous chromosomes occurs during meiosis and results in the exchange of chromosomal regions. The probability of a recombination event occurring between two loci is proportional to the distance between them. Thus, the frequency of recombination between two loci provides an estimate of the distance separating them. Positioning of a third locus can be achieved by determining the frequency of recombination which occurs between it and the other two loci. The stepwise analysis of many genetic loci in this manner can be used to generate a linkage map of each chromosome. *See* RECOMBINATION (GENETICS).

Accurate determination of linkage between two loci requires the analysis of large numbers of individuals from multigeneration families for recombination events. In human populations these are seldom available, and it is thus necessary to combine the information obtained from many smaller families by using so-called lod (log odds) scores. The lod score gives the probability that a particular distribution and combination of traits would have occurred if the two loci were linked with a particular recombination fraction (σ), and is defined by the following equation:

$$\text{lod} = \log_{10} \frac{\begin{array}{c}\text{probability of linkage between}\\ \text{two loci assuming } \sigma = \sigma\end{array}}{\begin{array}{c}\text{probability of the two loci}\\ \text{being unlinked } (\sigma = 0.5)\end{array}}$$

Lod scores from individual families can be summed to give a cumulative score. A lod score of 3.0 is considered good evidence for linkage, and one of -2.0 is indicative of its absence.

The observation of recombination events is dependent upon the possibility of distinguishing between equivalent loci on homologous chromosomes. Only loci with a minimum of two alleles can be used in linkage analysis, and these alleles must be present in the population at a reasonable frequency to be informative. Although some genes, such as those for blood groups, have suitable allelic frequencies, their number is limited.

Restriction-fragment-length polymorphism markers. Polymorphisms (genetic variations among individuals in a population) at the DNA sequence level, called restriction-fragment-length polymorphisms, have provided an unlimited source of codominantly inherited markers ("landmarks") ideal for the construction of linkage maps. Restriction enzymes, isolated from bacteria, recognize and hydrolyze DNA at specific 4–8-bp sequences. More than 200 restriction enzymes are known, each of which recognizes a particular sequence. Any sequence change within a restriction enzyme recognition site prevents cleavage there. In the absence of such a site, cleavage with the enzyme produces a larger DNA fragment. This absence can be detected by Southern blotting (**Fig. 1**), in which genomic DNA is cleaved by the restriction enzyme and fractionated according to size

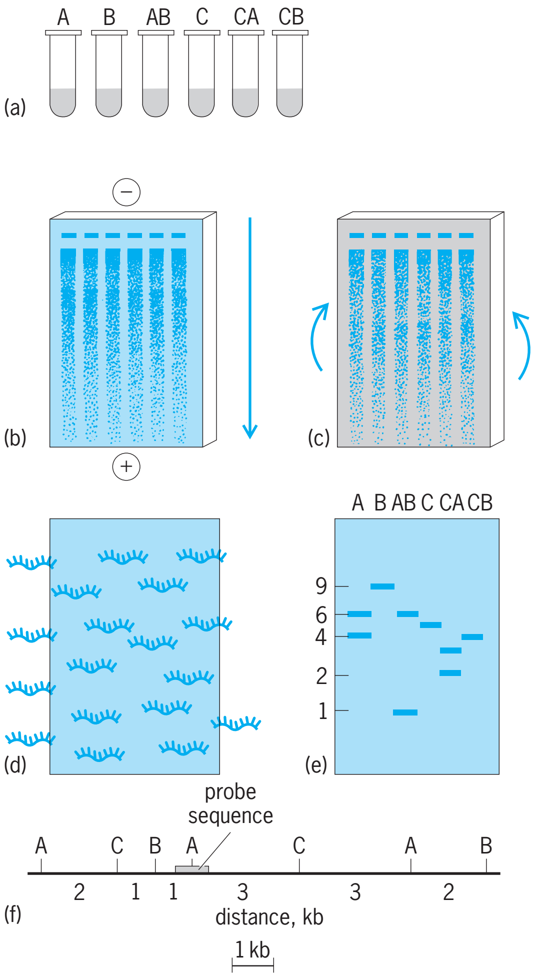


Fig. 1. Southern blotting. (*a*) Genomic DNA is digested singly and doubly with restriction enzymes A, B, and C. (*b*) Digested DNA is fractionated by size-dependent migration through a gel matrix in an electric field. (*c*) Denatured DNA is transferred and immobilized onto a nylon membrane. (*d*) Filter is incubated with a radiolabeled single-stranded probe in aqueous solution. (*e*) Duplexes formed between the single-stranded probe and genomic DNA are detected by autoradiography, showing the calculated size of DNA fragments. (*f*) Comparison of the products obtained from single and double digestion allows the construction of a restriction map (1 kb = 1000 base pairs).

by electrophoresis through a gel matrix. The DNA is transferred to a membrane and probed with a DNA segment near the polymorphic site (**Fig. 2**). *See* RESTRICTION ENZYME.

Short tandem repeat (microsatellite) markers. The development of short tandem repeat markers, or microsatellites, spanning the genome greatly facilitated human linkage analysis, and the era of positional cloning came to the forefront. Short tandem repeats are repetitive DNA sequences of two to five nucleotides that are scattered throughout the genome in abundance. They are often used as genetic markers because they tend to vary between individuals and can be easily typed using a conjuction of PCR assays and get electrophoresis. One of the first efforts of the Human Genome Project was to identify and map at least one Short Tandem Repeat marker per centimorgan (cM). Currently, more than ten thousand Short Tandem Repeats have been mapped at high density to produce genetic maps accessible through a web interface. *See* POLYMORPHISM (GENETICS).

Applications. Linkage maps provide the starting point for gene mapping. By establishing linkage between a gene and known polymorphic markers, the position of that gene can be determined. That technique is particularly useful for mapping of genes without a known biochemical function, such as those involved in diseases. Furthermore, closely linked markers can be used for prenatal diagnosis in families with a hereditary disease. The cosegregation of a linked marker with a disease gene permits the prediction of whether an offspring has the disease gene simply by following the inheritance of the marker. Linkage analysis also permits the simultaneous mapping of multifactorial traits to many chromosomal regions. *See* HUMAN GENETICS.

Association studies. Linkage analysis has been very effective in identifying rare, single-gene disorders (or Mendelian disorders) but has been less successful at locating genes responsible for common complex diseases, such as cardiovascular disease or cancer. In common diseases, each gene variant contributes only modestly to disease risk, and thus detection of these small effects is difficult unless the number of families studied is very high.

A complementary approach to linkage mapping is to identify disease genes by testing for an association between a specific polymorphic variant and the disease, that is, to conduct an association study. This is done by comparing the frequency of the specific polymorphism in a group of affected individuals with the frequency of the selected polymorphism in a group of unaffected controls.

There are two approaches to association studies: the "direct approach," in which each putative causal variant is tested for correlation with the disease and the "indirect approach," in which the polymorphic variants are used as genetic markers to detect the association between a particular genomic region and the disease.

In order to provide the tools needed to conduct association studies in an efficient and cost-effective manner, the international HapMap project was created.

The aim of this project is to determine the common patterns of DNA sequence variation in the human genome by the characterization of sequence variants, their frequencies, and the correlations between them in DNA samples from populations of

Fig. 2. DNA sequence polymorphisms. Restriction-fragment-length polymorphisms. A polymorphic restriction site R* is detected by the probe. A 3-kb fragment is produced when the site is absent, and a 2-kb fragment when it is present.

African, Asian, and European descent. The tools generated will allow the application of the indirect association approach to any functional candidate gene, any region suggested by linkage studies, or ultimately to whole genome scans searching for disease risk factors.

The polymorphic DNA markers of choice for association studies are single nucleotide polymorphisms. Single nucleotide polymorphisms are the most abundant DNA variation in humans (occurring on average in 1/300 bases), and they are caused by a single base change in the genome sequence. They are found both in genes and noncoding sequences, and some of them are expected to be the cause for common diseases.

A variety of single nucleotide polymorphism genotyping technologies have been developed recently. They vary in their assay formats, detection methods, and throughput capabilities. Some of the methods are capable of generating from several thousand to well over one million single nucleotide polymorphism genotypes per day, making the analysis of whole genomes possible. These technologies combine robotics, multiplex assays, and specialized software to deliver high-confidence and highly accurate genotypes.

DNA sequencing. Recombinant DNA technology has led to the mapping and analysis of genes at a molecular level. Fundamental to this has been the ability to produce large quantities of highly purified DNA sequences for further analysis by using molecular cloning techniques. Restriction enzymes can be used to insert DNA sequences into vectors propagated in bacterial hosts. Plasmid vectors consist of small, self-replicating circles of DNA, which can be used to clone segments of DNA of up to 10,000 bp (10 kilobases, or kb). Larger segments can be propagated in vectors derived from bacteriophage λ. Part of the λ genome is not essential for lytic growth and can be replaced with up to 20 kb of exogenous DNA. This recombinant-DNA molecule is incorporated into phage particles in the test tube and used in multiple rounds of infection of *Escherichia coli* cells. The DNA segments of 45 kb can be cloned into cosmid vectors, which contain only the sequences from bacteriophage λ necessary for incorporation in phage particles. However, these can be used only for a single round of infection and are often unstable.

In order to clone large genomic regions, two main cloning systems were developed: yeast artificial chromosomes and bacterial artificial chromosomes. Yeast artificial chromosomes clones contain DNA sequence inserts of more than 1 Mb (megabase) in size. They were used to constuct the first human and mouse physical maps. However, yeast artificial chromosomes are not very stable and have a tendency to rearrange their insert content. Bacterial artificial chromosomes clones have smaller DNA sequence inserts (100–200 kilobases), but they are more stable and have proven extremely useful for constructing second-generation physical maps of large genomes. Bacterial artificial chromosomes became the cloning vector of choice for the mapping of the human genome and the mapping of the genomes of the model organisms that followed.

The order of nucleotides in a DNA fragment can be determined by DNA sequencing. In the Sanger method (also known as the dideoxy chain termination sequencing method), a complementary sequence to a single-stranded template is synthesized by the enzyme DNA polymerase. Synthesis of the complementary strand is terminated at specific residues by the incorporation of a modified nucleotide. A separate reaction is performed for each nucleotide and results in a family of molecules, each of which terminates at that nucleotide. These products can be fractionated on the basis of size by electrophoresis through an acrylamide gel, and the sequence can be read from a ladder of fragments of different size (**Fig. 3**). *See* ELECTROPHORESIS.

During the sequencing efforts of the Human Genome Project, crucial enhancements were made toward increasing the accuracy and efficiency of the Sanger method. Advances in laser instrumentation, allowing for better automated detection of fluorescently labeld DNA molecules; the development of capillary-based sequencing instruments capable of analyzing 96 samples in parallel; and improvements

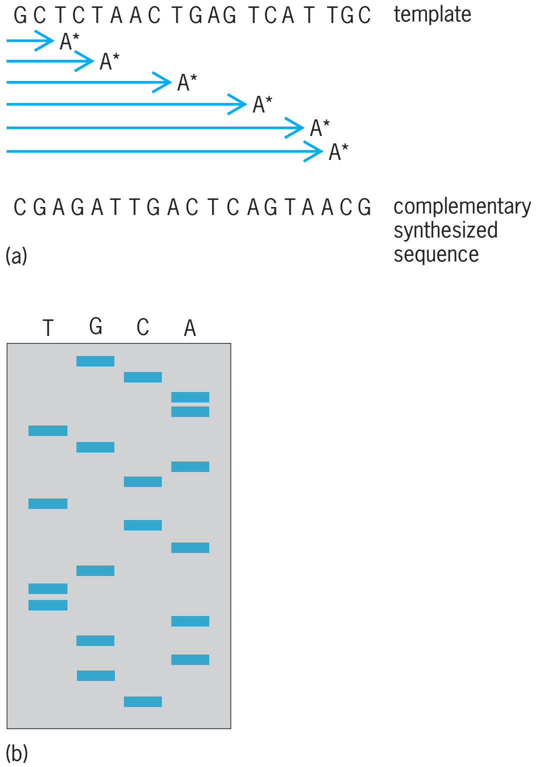


Fig. 3. DNA sequencing. (*a*) Complementary strand of DNA is synthesized to a single-stranded template by the enzyme DNA polymerase. Chain growth is terminated upon the incorporation of a modified nucleotide, in this case A*, which results in a family of fragments each of which terminates in the residue A. A similar reaction is performed for each of the other three nucleotides, T, G, and C. (*b*) The products of each reaction are fractionated according to size by electrophoresis. The order of nucleotides is read from the bottom of the gel (the smallest fragment) to the top.

in the biochemistry of the reaction made possible the dramatic increase in throughput (500–800 quality bases per reaction; 1.5 million bases per 24 h) and at least a 100-fold reduction in costs. In addition, software systems were developed for the analysis and quality assessment of the primary sequence data, and viewers were created for inspection and editing of the resulting sequence assemblies.

DNA sequencing continues to evolve toward faster, more affordable, and more sensitive technologies. In particular, nanotechnology approaches—capable of a high level of miniaturization, automation, parallelization, and integration—are making progress toward this end. These improvements will not only allow sequencing of entire new genomes at a reasonable cost but will also open the possibility of sequencing additional human genomes. The ability to decode the genomic sequence of virtually any organism will provide unparalleled access to the fundamentals of biology. The ability of resequencing the human genome without restrictions will have an enormous impact on understanding human genetic variability and the current approaches to manage and cure disease.

Sequence mapping. The expression of a gene is dependent upon the synthesis of a messenger RNA (mRNA) molecule and its translation into the amino acid sequence of a protein by ribosomes in the cytoplasm. The DNA copies of mRNA (complementary DNA, or cDNA) from cells are synthesized by using the viral enzyme reverse transcriptase, and cloned for further analysis. Comparison of cDNAs with genomic sequences reveals that the DNA encoding many eukaryotic genes is discontinuous. The blocks of coding sequences (exons) of a gene are brought together in the mRNA molecule by the splicing out of intervening sequences (introns). The presence of introns results in genes being distributed over a much larger stretch of DNA than would be expected from the analysis of their mRNA or protein sequences. The size and number of exons and introns vary widely between genes. For instance, the β-globin gene contains only two introns of 14 and 331 bp and encodes a 620-bp mRNA. In comparison, the gene disrupted in Duchenne's muscular dystrophy encodes a 16-kb mRNA that comprises at least 60 exons and extends over 2×10^6 bp of DNA. *See* EXON; INTRON; RIBONUCLEIC ACID (RNA).

Before the Human Genome Project, mapping of the intron-exon structure of a gene was approached by using restriction enzymes. Restriction sites present in the genomic DNA but absent in the cDNA indicate the presence of some introns. However, precise mapping of the boundaries of the intron-exon structure can be achieved only by obtaining the genomic DNA sequence.

In the past, this required the isolation of large sections of genomic DNA by chromosome-walking techniques. A library for this is constructed by partially digesting genomic DNA with a frequently cutting restriction enzyme. This produces a series of overlapping DNA fragments that cover the entire genome. These are cloned into bacteriophage or cosmid vectors, and the resulting library is screened by hybridization with a radioactive probe. Positively reacting clones that contain genomic DNA surrounding the probe are purified and restriction-mapped. New probes are isolated from the ends of these clones and used to obtain a second set of clones that overlap the original clone. Repetition of this process results in the isolation of a series of clones that contain an ordered set of overlapping genomic fragments. *See* BACTERIOPHAGE; GENETIC ENGINEERING.

The isolation of genomic sequences surrounding a gene by chromosome walking may result in the mapping of gene clusters. For example, the human β-globin gene family stretches over 50 kb on chromosome 11. It contains five related functional genes that are active at different stages in human development. All of the genes are composed of three exons and probably arose from a single ancestral gene. Duplication of the ancestral gene by unequal crossing-over followed by sequence divergence over time has resulted in the evolution of five related but distinct genes. A homologous but functionally inactive β-globin pseudogene was also mapped within this region. Pseudogenes are genes that were presumably once active but have become defunct from the accumulation of mutations. The comparison of sequences surrounding genes has led to the identification of sequences involved in the regulation of genes. *See* PARASEXUAL CYCLE.

With the map and sequence of the human genome available, most of the steps described above can be circumvented, and the mRNA sequence can be compared with the genomic sequence without any additional experimental work. The exon-intron boundaries can be defined using sequence homology software, which compares the mRNA sequence to the publicly available human genome sequence.

Mapping molecular mutations. Many mutations are the result of small, subtle changes within the DNA sequence of a gene. The substitution of a single base pair within the reading frame of a gene may result in the incorporation of a different amino acid in the protein or premature termination of mRNA translation. These changes can eliminate or drastically alter the activity of a protein. For example, the reduced affinity of hemoglobin for oxygen in sickle cell anemia is caused by a single amino acid substitution resulting from a single base change in the β-globin gene. These mutations, as well as others, are mapped by comparing the DNA sequence of the mutant gene with that of the wild type. The sequence of specific regions of the human genome in large populations can be ascertained by the selective amplification of those regions by using PCR. Oligonucleotides flanking the target sequence are used as primers for repetitive cycles of DNA synthesis of that sequence in the test tube (**Fig. 4**). The resulting amplified product can be directly subjected to DNA sequence analysis. Alternatively, the presence or absence of a known polymorphic restriction site within the product can be tested directly by enzyme digestion

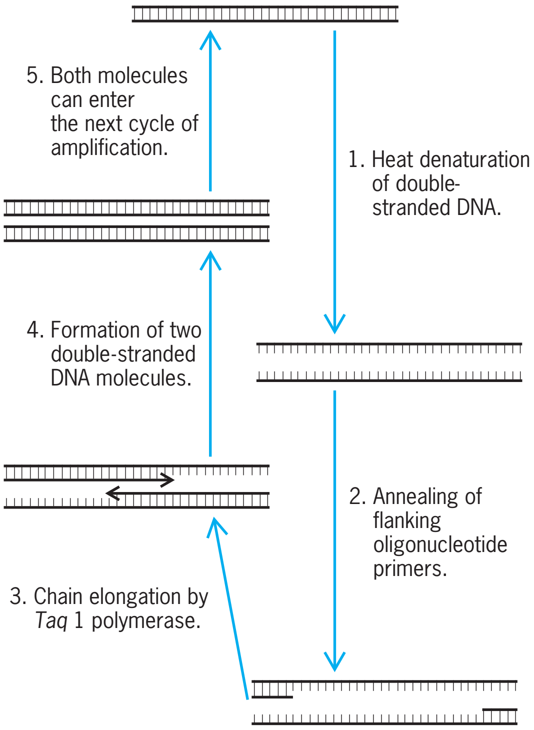


Fig. 4. Polymerase chain reaction. Regions of DNA defined by flanking oligonucleotides are amplified by successive cycles of denaturation, primer annealing, and chain elongation. Ideally, each round of amplification would double the target sequence. The resulting product can be analyzed with restriction enzymes and by DNA sequencing.

and gel electrophoresis. This simplifies the typing of individuals for restriction-fragment-length polymorphisms in linkage analysis. Furthermore, the magnitude of DNA amplification by polymerase chain reaction permits the analysis of single sperm cells. Each sperm cell is the product of a single meiosis. The amount of recombination occurring between two polymorphic loci can be accurately measured by simultaneously amplifying and typing those loci in a large number of individual sperm cells. Although this analysis is restricted to DNA sequence polymorphisms, it may provide a much simpler and more accurate method of measuring the genetic distance separating such loci than is offered by extensive family studies. *See* GENE AMPLIFICATION; MEIOSIS.

Reverse genetics (or positional cloning). Traditionally, the mapping and isolation of genes was dependent upon prior knowledge of the activity or amino acid sequence of the respective protein products. Before 1980, very few human disease genes had been identified, because for the majority of genes, including those involved in genetic disorders, this information was not available. Such genes are isolated by reverse genetics, also called positional cloning. This methodology identifies disease genes based solely on their approximate chromosomal location. The first stage in positional cloning involves linkage mapping of the disease to the smallest possible region of the genome in affected families. Next novel candidate genes are identified in the region of linkage. The final step is to show that patients have mutations in a gene that are not present in control subjects. Before these steps could be completed, researchers had to produce high-resolution genetic maps and build clone contigs in order to identify and sequence the genes of interest. This process is now easier to implement due to the resources—maps, clones, sequence, expression and phenotypic data—provided by the Human Genome Project as well as other genome research projects. Thus, the experimental part is reduced to linkage mapping of the disease gene to a region of the genome and to the selection and screening of the candidate genes for mutations in patients.

Candidate genes can be identified by the presence of unmethylated CpG-rich stretches of DNA, the conservation of DNA sequence with other animal species, or the identification of regions that are transcribed into mRNA. The protein sequence of these genes can be obtained by the isolation and sequence analysis of their cDNAs, providing a starting point for examining the gene function. Often the disease gene can be identified by its disruption of activity in affected individuals. Many genes, including those involved in Duchenne's muscular dystrophy, sex determination, and chronic granulomatosis disease, have been mapped and isolated by reverse genetics. *See* GENETIC CODE; MOLECULAR BIOLOGY.

Matthijs Smith; Leroy Hood

Bibliography. R. DeSalle and M. Yudell, *The Genomic Revolution: Unveiling the Unity of Life*, Joseph Henry Press, 2002; B. Lewin, *Genes*, 2004; G. S. Stent and R. Calender, *Molecular Genetics: An Introductory Narrative*, 1978; R. K. Saiki et al., Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase, *Science*, 239:487–491, 1988; T. Strachon and A. P. Read, *Human Molecular Genetics*, 3d ed., 2003; J. D. Watson et al., *The Molecular Biology of the Gene*, 5th ed., 2000; T. J. White, N. Arnheim, and H. A. Erlich, The polymerase chain reaction, *Trends Genet.*, 5:185–194, 1989.

Genetics

The science of biological inheritance, that is, the causes of the resemblances and differences among related individuals.

Genetics occupies a central position in biology, for essentially the same principles apply to all animals and plants, and understanding of inheritance is basic for the study of evolution and for the improvement of cultivated plants and domestic animals. It has also been found that genetics has much to contribute to the study of embryology, biochemistry, pathology, anthropology, and other subjects. *See* BIOCHEMISTRY; EMBRYOLOGY; PATHOLOGY; PHYSICAL ANTHROPOLOGY.

Genetics may also be defined as the science that deals with the nature and behavior of the genes, the fundamental hereditary units. From this point of view, evolution is seen as the study of changes in the gene composition of populations, whereas embryology is the study of the effects of the genes on the development of the organism. *See* POPULATION GENETICS.

Study and analysis. The geneticist uses various methods of study, nearly all of which rest on the presence of differences between individuals. These techniques do not make it possible to analyze the portions of the hereditary makeup of individuals for which differences cannot be found. If differences occur, it is also necessary that they be transmitted to later generations; and in general the full analysis will also require the presence in the material of sexual reproduction or some analogous process that allows the recombination of inherited properties from different individuals. In brief, what the geneticist usually does is to cross diverse individuals and study the descendants (or in such material as humans, analyze existing pedigrees). Such a study must usually be carried through at least two successive generations and requires enough individuals to establish ratios between the classes present, since conclusions are usually based on these. The mechanism of heredity is such that its analysis requires the use of probability theory, although for much of the subject only very simple statistical concepts are needed. *See* BIOMETRICS; BREEDING (ANIMAL); BREEDING (PLANT); CHROMOSOME; CHROMOSOME ABERRATION; HETEROSIS; HUMAN GENETICS; MENDELISM. A. H. Sturtevant

Molecular genetics. A major discipline within the field of genetics that describes the basis of inheritance at the molecular level, molecular genetics had its beginnings in the 1940s. At that time, work on simple, single-celled organisms revealed that the genetic material in most organisms is deoxyribonucleic acid (DNA), a very long polymer made up of four different nucleotide monomers: adenosine (A), guanosine (G), cytosine (C), and thymidine (T). A landmark discovery was the demonstration in 1953 by James Watson and Francis Crick that the genetic material consists of two DNA strands wrapped around each other to form a double helix. The two strands are held together by specific noncovalent, weak interactions between the nucleotides: A always binds to T, and G always binds to C. These long, double-stranded DNA molecules make up chromosomes, which are divided into thousands of different functional units called genes. The genes specify the structure and function of an organism. *See* ALLELE; DEOXYRIBONUCLEIC ACID (DNA); GENE; NUCLEIC ACID.

The field of molecular genetics focuses on two general questions: how do genes specify the structure and function of organisms, and how are genes replicated and transmitted to successive generations? Both questions have been answered. Genes specify organismal structure and function according to a process described by the central dogma of molecular biology: DNA is made into ribonucleic acid (RNA), which specifies the structure of a protein, a polymer of 20 different amino acids that carries out a particular function. In the first step of gene expression, DNA is transcribed into messenger RNA (mRNA)—a single-stranded polymer of nucleotides very similar to DNA—by the enzyme RNA polymerase. In the second step, the mRNA molecule serves as a template for protein synthesis, which is carried out by a complex machinery that comprises a particle called a ribosome and special adapter RNA molecules called transfer RNA. *See* RIBONUCLEIC ACID (RNA); RIBOSOMES.

The specific sequence of nucleotides in the DNA (A, T, G, C) specifies gene function in two ways. First, particular sequences signal the RNA polymerase to begin and end synthesis of the mRNA. In addition, specific DNA sequences ensure that RNA polymerase transcribes certain genes only in the correct cell type at the proper time. A major goal of molecular genetics is to learn how DNA sequence determines the regulation of gene expression. The second way in which the nucleotide sequence specifies gene function is by determining the structure of the encoded protein. The sequence of the mRNA is read by the ribosome in groups of three nucleotides, each of which is called a codon. The four different nucleotides are arranged in 64 codons that specify the 20 amino acids in proteins. Therefore, the DNA sequence determines the structure and function of an organism. Molecular geneticists have learned how to determine large amounts of DNA sequence with relative ease. Large-scale projects have been undertaken to determine the complete DNA sequence of several organisms, including humans. Although it is clear how the sequence of DNA determines the sequence of protein, a major challenge remaining for molecular geneticists is to explain how the sequence of a protein determines its particular function. *See* GENETIC MAPPING.

The structure of DNA provides a simple mechanism for genes to be faithfully reproduced: the specific interaction between the nucleotides means that each strand of the double helix carries the information for producing the other strand. That transfer is accomplished by several enzymes, including DNA polymerase. It copies each DNA strand into two molecules of double-stranded DNA that are apportioned to each new cell during cell division. Rarely, DNA polymerase makes a mistake during replication, or the DNA becomes damaged, and an altered gene results. Such mutations can change the function of the encoded protein and thereby lead to evolutionary change. In addition, DNA molecules can recombine through complex processes that involve the breakage and rejoining of two molecules and lead to different gene combinations that can alter the structure and function of the organism. *See* GENETIC CODE; GENETIC ENGINEERING; MOLECULAR BIOLOGY; MUTATION. Mark Johnston

Bibliography. M. Levitan, *Textbook of Human Genetics*, 3d ed., 1988; B. Lewin, *Genes*, 3d ed., 1987; G. M. Malacinski and C. Simonsen, *Molecular Genetics of Mammalian Cells: A Primer in Development Biology*, 1986; G. S. Stent and R. Calender, *Molecular Genetics*: *An Introductory Narrative*, 1978; J. D. Watson and N. H. Hopkins, *Molecular Biology of the Gene*, 4th ed., 1988.

Genomics

The comprehensive analysis of an organism's genome (the full complement of inherited information-bearing genetic material residing in an organism's chromosomes), including, but not limited to, its genes and the signaling codes regulating gene expression. As a discipline, genomics focuses on an organism's entire set of genes rather than on individual genes. The term was coined in 1986 to name a journal oriented to the mapping, sequencing, and analysis of genomes.

Typically, genetic material is synonymous with DNA, although some viruses use RNA rather than DNA to encode genomic information. DNA is a polymer consisting of four nucleotide building blocks, also referred to as bases, called A (adenine), C (cytosine), G (guanine), and T (thymine). During the 1950s and 1960s, seminal advances in understanding DNA's double-helical structure, its mechanism of replication, and its code for specifying the amino acid content of proteins led to a significant insight: "Sequence"—the arrangement of nucleotides (bases) in some particular order along a DNA strand—is the key to deciphering a genome's information content (**Fig. 1**). This means that the information in the genome is digital; that is, it consists of strings of discrete letters whose order provides meaning or information just as the 0s and 1s in computer programs provide meaning. The field of genomics, therefore, is directed to the informative features embedded in a given genome's DNA sequence. In practice, a global genomic analysis requires an empirically determined DNA sequence called a "reference genome" for a species of interest. The best-known example of a comprehensive DNA sequence-based genomic analysis is the Human Genome Project. It is fair to say that no other single project has so transformed the practice of both biology and medicine. Owing to spectacular advances in DNA sequencing technology resulting from the Human Genome Project, genomes from hundreds, soon to be thousands, of species of medical, agricultural, industrial, and phylogenetic significance have been made available for biologists to explore.

Importance. Genomic information is crucial to understanding biology. In essence, a species' genes and regulatory circuitry determine possibilities for, and set constraints on, the structure and function of its individual members. Genomic programming allows an acorn to develop into an oak tree and a caterpillar into a butterfly. An acorn will not, and cannot, produce a butterfly, however, because there are significant differences in specialized genes and regulatory circuitry between plant and insect genomes. Genomic information allows biologists to catalogue and compare "parts lists" for different species. Parts lists might include DNA sequences identified as protein-coding and regulatory RNA genes. In addition, there are short sequence stretches called promoters, enhancers, and silencers that, with the help of proteins called "transcription factors," control whether, when, and how much a particular gene is transcribed. Parts lists also contain sequences that serve as origins of DNA replication and otherwise affect the functioning of chromosomes. In essence, differences within and between species at the nucleotide sequence level provide the informational foundation for a detailed molecular understanding of cellular function and organism development.

Scientific approach. Genomics aims to develop large-scale, high-throughput, cost-effective strategies for the analysis of genomes. These strategies are used first to decipher what a given genome's DNA sequence is and then to determine or predict what information content that sequence contains. Thus,

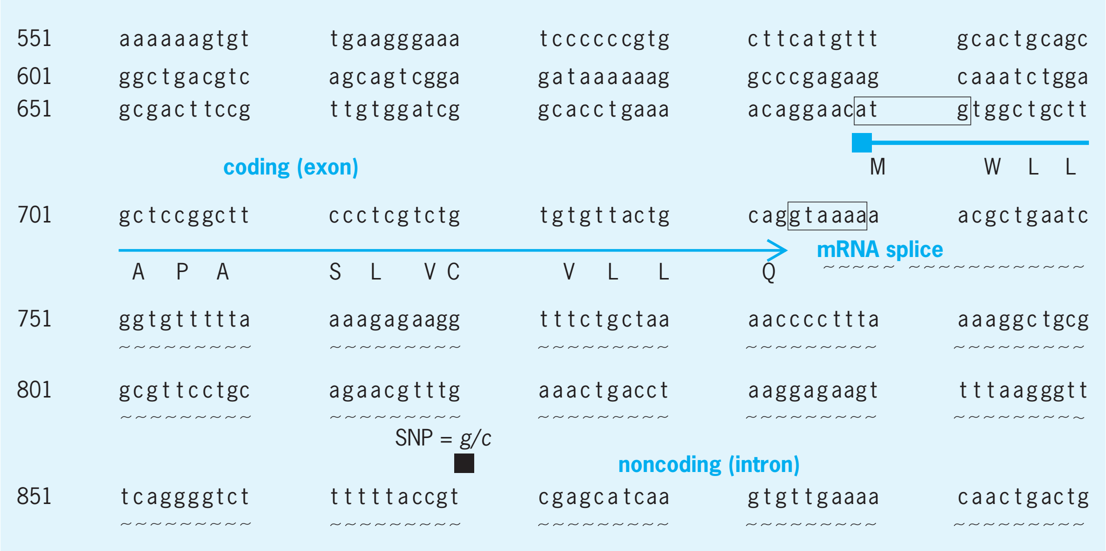


Fig. 1. Informative features within the DNA sequence. This stretch of sequence represents a 350-base portion of a pufferfish toll-like receptor gene containing the first exon and a portion of the first intron. The black square depicts the location of a sequence variation (polymorphism). The color arrow indicates the protein-coding sequence, with the amino acids depicted in capital letters.

the methodologies of genomics encompass two distinct activities:

1. *Sequence determination*: Using best-practice technologies to delineate the order of nucleotide building blocks in an organism's chromosomes or genome.

2. *Sequence interpretation*: Using computational and wet-laboratory procedures to extract meaningful biological information from a genomic sequence or sequences.

A detailed and comprehensive analysis of a species' genome requires an accurately assembled reference sequence for each of the chromosomes. Once the reference sequence has been determined, the location and frequency of polymorphic variations can be assessed. "Polymorphism" is the name given to the small changes in DNA sequence that distinguish a genome of one member of a species from the genomes of other members. Polymorphisms are identified by sequencing genomes, or portions of genomes, of two or more members of the species. From a technical standpoint, obtaining the initial reference sequence, called "de novo sequencing," is considerably harder than selectively "resequencing" genomes for discovering polymorphisms.

Ideally, a reference genome sequence will be of "finished" quality. In a finished genome, the accuracy of each base is greater than 99.99%. Perhaps more importantly, the sequence consists of gap-free contiguous strings of correctly ordered As, Cs, Gs, and Ts that correctly represent the organism's chromosomes. Several eukaryotic genomes have been finished—for example, human, fruitfly, worm, and baker's yeast—and hundreds of microbial genomes have been completely sequenced. The ease of obtaining a finished sequence for any given genome is a function of its size, complexity, and sequence composition. Some genomes are harder to finish than others because they contain multiple gene duplications, extremely high or low GC content, or regions that are difficult to propagate with the commonly used cloning vectors (agents into which small fragments of DNA are inserted for amplification prior to sequencing). Unfinished genomes are provided in "draft" form. Although their quality is lower than finished sequences, draft sequences nonetheless provide valuable information about a species' genomic content.

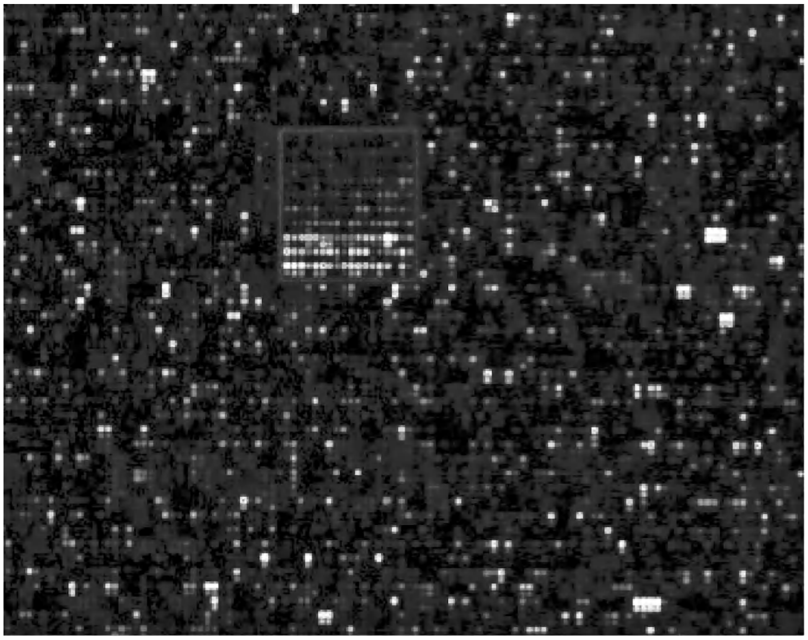


Fig. 2. Image of an Affymetrix U133 Plus 2.0 Genechip hybridized with biotin-labeled target derived from prostate luminal cells. The array contains probe sets that detect 47,000 transcripts that represents the human transcriptome. The intensity of signal in the image represents the abundance of the mRNA which the probe set targets. Signal intensity is depicted as a gradient from gray to white, where gray is the lowest signal and white is the greatest signal. (*Asa Oudes, Institute for Systems Biology*).

Once a reference genome sequence has been obtained, a variety of computational and experimental strategies are employed to identify its significant features. Analyses include but are not limited to the following types of information:

1. *Genomic landscape features*: Delineating gene models, order of genes along a chromosome, gene duplications, common repeats, hot and cold spots of recombination, and features related to chromosome organization and function.

2. *Regulatory signals*: Identifying sequence motifs that affect the occurrence and extent of transcription of specific genes or groups of genes under varying cellular conditions.

3. *Cross-species comparisons*: Identifying blocks of synteny (used here to mean regions where the order of the genes is approximately the same in the genomes or two or more species) using other reference genomes, relating a genome's information content to that of other species, and discovering highly conserved blocks of noncoding sequence that may be indicative of new and important types and functions of sequence information.

4. *Sequence variations*: Applying resequencing technologies to genomes of individuals within a species to catalogue single nucleotide substitutions, insertion-deletion polymorphisms, gene-copy-number polymorphisms, and chromosomal rearrangements.

5. *Disease association studies*: Identifying candidate genes whose malfunctions cause deleterious effects.

A large-scale international effort, called the ENCODE project (Encyclopedia of DNA elements) is under way to characterize and identify functional features of the human genome sequence. It is anticipated that approaching the problem from a variety of experimental and computational perspectives will illuminate the information in the human genome in ways that cannot currently be foreseen.

Genomics and systems biology. Genomics has stimulated the development of numerous systems biology-related technologies aimed at examining gene expression profiles, gene/protein interactions, and regulatory networks. These high-throughput technologies perform multiple measurements simultaneously so that interactions can be assessed in a comprehensive manner. For example, microarrays take advantage of the ability of RNA to hybridize to known stretches of complementary DNA placed in defined locations on a chip. Microarrays are used to compare, under varying conditions, the increases and decreases in expression of thousands of gene transcripts (mRNAs) in parallel. Transcripts from samples of cells or tissues can also be identified and assessed quantitatively based on high-throughput parallized measurements (that is, measurements made in parallel) of short, unique signature sequences. DNA microarrays are also being used to identify regions of sequence that are bound by transcription factor proteins under experimentally specified conditions. The results reveal sequence signals important to regulating biological functions or disease pathology (**Fig. 2**). Application of these technologies and others in a concerted fashion is illuminating genetic, metabolic, signaling, and regulatory networks. These networks govern a cell's responsiveness to its internal developmental programming and external environmental perturbations. Studies of such networks are termed systems biology.

Moving forward. Advances in sequencing technology have vastly increased throughput since the Human Genome Project was begun in 1990. Significant improvements in base-call accuracy, read length, automation, machine capacity, instrument run time, and assembly algorithms have made de novo genomic sequencing both cheaper and easier to do. However, in terms of future needs, enormous reductions in cost are required. To this end, several "new generation" sequencing technologies based on highly parallel sequencing (millions of samples at one time) or single molecule sequencing are being developed. These aim to improve the efficiency and throughput of de novo sequencing, polymorphism detection, and transcriptional profiling while greatly decreasing cost and making affordable personal human genome sequences a possibility. Individual human genome sequences will be a key for predictive medicine of the future.

Inspired by this optimism, an ambitious project surpassing the scale of the Human Genome Project was launched at the beginning of 2006. This effort, named "The Cancer Genome Atlas" (TCGA), has the goal of identifying key mutational events responsible for oncogenesis and metastasis in various types of cancerous cells. Because cancerous cells accumulate thousands of mutations during their rounds of propagation, it is a daunting task to identify those key changes that account for differences in malignancy among the different cancer types. However, a more detailed molecular and systems understanding of how cellular networks are perturbed in various types of cancer will undoubtably lead to more effective drug targets and therapies.

As a logical extension of The Human Genome Project, intensive efforts are under way to construct detailed "haplotype maps." These maps delineate polymorphic variations that group together in "blocks or haplotypes"of sequence shared by differing human subpopulations. Advances in technologies for identifying polymorphisms and genotyping individuals have promoted an increased interest in "personalized medicine"—the right treatment for the right patient at the right time. Personalized medicine relies on the ability to coordinately correlate the sequence polymorphisms of multiple genes with specific effects, such as susceptibilities to diseases, the efficacy of pharmacological interventions, and the severity of side effects.

The explosion of sequence information produced by genomic technologies demands progress on two additional fronts: development of statistical algorithms and computational modeling tools for interpreting the information; and examination of the social, legal, and ethical implications of genomic research. *See* DEOXYRIBONUCLEIC ACID (DNA); GENE; GENE ACTION; GENE AMPLIFICATION; GENETIC CODE; HUMAN GENOME PROJECT; MOLECULAR BIOLOGY.

Lee Rowen; Leroy Hood

Bibliography. L. Gonick and M. Wheelis, *The Cartoon Guide to Genetics*, rev. ed., Harper-Collins, New York, 1983, updated 1991; L. Hood et al., *Biological Information and the Emergence of Systems Biology*, Roberts and Company, 2006; The sequence of the human genome, *Science*, 291, no. 5507, Feb. 16, 2001; G. Smith, *The Genomics Age: How DNA technology Is Transforming the Way We Live and Who We Are*, Amacom, New York, 2005; Initial sequencing and analysis of the human genome, *Nature*, 409:860–921, 2001.

Gentianales

An order of flowering plants (angiosperms) in the euasterid I group of the asterid eudicotyledons. The order consists of five families and approximately 17,500 species. The circumscription of Gentianales has been widened to include Rubiaceae, largely on the basis of deoxyribonucleic acid (DNA) sequence data. Circumscription of Loganiaceae has also changed on the basis of these data, with many taxa, including the ornamental *Buddleja*, being transferred to Lamiales. The order is characterized by opposite leaves, frequent occurrence of alkaloids, and the presence of internal phloem in the wood (not in Rubiaceae).

Apocynaceae (including Asclepiadaceae), approximately 5000 species, are mostly tropical and subtropical and have a well-developed latex system, superior ovary, thickened and apically modified style, and carpels that are usually united only by the common style or stigma. *Asclepias* (milkweed), *Catharanthus* (rosy periwinkle, the source of the powerful anticancer drugs vincristine and vinblastine), *Hoya* (wax flower), *Nerium* (oleander), and *Stapelia* (carrion flower) are well-known genera. Rubiaceae, approximately 11,000 species, are cosmopolitan, but particularly important in the tropics. They have stipulate leaves and an inferior ovary. The widespread occurrence of alkaloids makes the family economically significant; the major products are coffee (*Coffea*) and the antimalarial drug quinine (*Cinchona*). Other genera, including *Gardenia*, *Ixora*, and *Pentas*, are widely used in horticulture. Gentianaceae, approximately 1200 species, are principally temperate and subtropical. They often contain bitter iridoid compounds, and some are used medicinally or in bitter alcoholic beverages (such as *Gentiana*, gentians). Several genera are widely grown ornamentals, including *Exacum* (Arabian violet), *Gentiana*, and *Lisianthus* (prairie gentian). The remaining two families are relatively small but contain some genera of economic importance. *Strychnos* (Loganiaceae) is the source of strychnine and the arrow poison curare, and Gelsemiaceae include the ornamental *Gelsemium* (allspice jasmine). *See* ASTERIDAE; MAGNOLIOPSIDA; PLANT KINGDOM; RAUWOLFIA; STROPHANTHUS; STRYCHNOS. Mark W. Chase