ASTRONOMY AND ASTROPHYSICS LIBRARY

Immo Appenzeller

High-Redshift Galaxies

Light from the Early Universe

A&A LIBRARY

Springer

ASTRONOMY AND ASTROPHYSICS LIBRARY

Immo Appenzeller

High-Redshift Galaxies

Light from the Early Universe

Springer

Prof. Dr. Immo Appenzeller
Landessternwarte Königstuhl
Königstuhl 12
69117 Heidelberg
Germany
i.appenzeller@lsw.uni-heidelberg.de

ISSN 0941-7834
ISBN 978-3-540-75823-5 e-ISBN 978-3-540-75824-2
DOI 10.1007/978-3-540-75824-2
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009926969

Cover design: eStudio Calamar S.L.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To Gisela

Preface

The high-redshift galaxies became a distinct research field during the final decade of the 20th century. At that time the Lyman-break technique made it possible to identify significant samples of such objects, and the new generation of 8 to 10-m telescopes resulted in first good spectroscopic data. Today the high-redshift galaxies have developed into one of the important topics of astrophysics, accounting for about 5–10% of the publications in the major scientific journals devoted to astronomy. Because high-redshift galaxies is a rapidly developing field and since new results are published constantly, writing a book on this topic is challenging. On the other hand, in view of the large amount of individual results now in the literature, and in view of the still growing interest in this topic, it appears worthwhile to summarize and evaluate the available data and to provide an introduction for those who wish to enter this field, or who, for various reasons, might be interested in its results.

The end of the first decade of the 21st century appears to be a good point in time to attempt such a summary. The current generation of ground-based 8 to 10-m IR-optical telescopes, the Hubble Space Telescope, and the most important large radio telescopes have by now been in operation since about one or two decades. Although these instruments will continue to produce important scientific results for some time to come, many of the initial programs exploiting their unique new possibilities have been completed. Further scientific progress with these facilities is expected to occur on a less rapid pace. Within about another decade a new generation of instruments (such as the James Web Space Telescope, the Atacama Large Millimeter Array, and new larger and innovative optical and radio telescopes) will take over the role as cutting-edge instrumentation. But these future instruments are still under construction or in advanced planning stages. Therefore, it seems timely and appropriate to take stock on what has been achieved during the past years, on what is missing, and on what has still to be done, and will be feasible in the future.

The main objective of this book is to summarize and to discuss our present knowledge of the galaxies with redshifts > 2. The light observed from these galaxies was emitted when the age of the universe was less than one quarter of the present value of 13.7 Gyr. At the epoch corresponding to a redshift of 2, galaxy formation had been completed and about a quarter of the cosmic star formation had already taken place, the cosmic star formation rate had attained its maximum, and the universe had qualitatively reached most of its present properties. The galaxies still

evolved significantly during the following $>$ 10 Gyr. But most of this evolution was of a quantitative nature.

In addition to summarizing and critically discussing the observational data and their present interpretation, the book gives a brief introduction to the observational methods, which make it possible to study these very distant and extremely faint radiation sources. This includes an assessment of the technical feasibilities and the physical limitations of the existing and of future ground-based and space-based instruments. The objective of these sections is to provide information and tools which could be of help in the planning of future observational programs and for the definition and development of new instrumentation.

As a result of major new astrophysical discoveries (such as dark matter, dark energy, and inflation) the distant universe has become a popular topic also among colleagues working in the field of fundamental physics and other related fields of science. The extreme conditions and processes in the early universe are often regarded as a particularly valuable test bench for emerging new physical concepts. The most distant galaxies were formed during the transition from the almost homogeneous very early universe to the structured universe in which we live today. Therefore, high-redshift galaxies contain important information on the physics and the properties of these early epochs.

In view of the relevance of observations of the high-redshift universe beyond astronomy, I have tried to make this book readable not only for astronomers, but for all scientists with a basic knowledge of physics. It is assumed that the reader is familiar with the basic concepts and results of modern physics, but reading the book does not require a prior knowledge of astronomy. For the readers without an astronomical background, the text includes brief descriptions of the basic concepts, terms, and notations used in the current astronomical literature. Furthermore, the first three chapters provide short overviews of our present knowledge on the main components of the local universe, and of the basic facts of cosmic evolution. The special terms, notations, and special units used in astronomy (which often appear puzzling to non-astronomers) are explained, when they appear for the first time.

These introductory sections may also be helpful for students of astronomy, who are interested in the present topic, but who have not yet completed all the basic astronomy courses. Naturally, the introductory sections can provide only short summaries of the wealth of information available in the basic astronomical textbooks. Moreover, only those facts and data are presented here, which are of interest in the context of high-redshift galaxies. Readers, who are interested in more comprehensive information on individual topics, will find references to the corresponding detailed literature.

In view of the high rate of papers published in the field of extragalactic astronomy, it is impossible to provide a complete list of references. Instead, the citations had to be restricted to the most important publications and to the most recent papers.

This book could not have been completed without the kind support of many colleagues who provided advice, unpublished information, figures, and data. Thanks are due to Ivan Baldry, Rychard Bouwens, Norbert Christlieb, Andrea Cimatti, Jos de Bruijne, Niv Drory, Daniel Eisenstein, Franz Elsner, Xiaohui Fan, Stephan Frank,

Armin Gabasch, Thomas Greif, Katarina Kovač, Mariska Kriek, Claus Leitherer, Dieter Lutz, Roberto Maiolino, Dörte Mehlert, Tohru Nagao, Stefan Noll, Masami Ouchi, Roderik Overzier, Will Percival, Max Pettini, Ryan Quadri, Sara Salimbeni, Alice Shapley, Robert Simcoe, Volker Springel, Otmar Stahl, Masayuki Tanaka, Christian Tapken, Christy Tremonti, Matteo Viel, Nolan Walborn, Christian Wolf, Makiko Yoshida, and Wei Zheng for providing figures or for allowing me to reproduce figures from their publications. I am particularly grateful to Jochen Heidt and to Christian Tapken for reading parts of the manuscript and for very helpful comments. Finally, I would like to thank Springer-Verlag for the efficient and pleasant cooperation during the preparation of this book.

Heidelberg, May 2009 *Immo Appenzeller*

Contents

Chapter 1
Introduction

1.1 The Relevance of Very Distant Galaxies

From observations of the Cosmic Microwave Background (e.g., [491]) and from other, independent astronomical observations we know that, averaged over sufficiently large scales, our universe is highly homogeneous and isotropic. This means that at a given instant of cosmic time our universe shows the same basic properties everywhere, regardless of the distance and regardless of the direction in which we observe. Therefore, at first glance, it may not appear worthwhile to observe distant objects. However, due to the finite velocity of light, a look at distant objects is automatically a look into the past. In astronomy we never see an object in its present state. We always observe the properties which the object had a long time ago, when the light was emitted. Since we know very well that our universe is evolving with time, observations of distant astronomical objects allow us to study the history of our cosmos. In the light of the most distant objects we directly see the universe in its earliest stages of development. Thus, distant objects are interesting not because of their distance in space, but because of their distance in time.

Although during the last decades our understanding of the large-scale properties of the universe has progressed significantly, we still do not know whether the cosmos has a finite volume or whether it extends to infinity. But we know that its extent is much larger than the volume and the distance which we can overlook (cf. Sect. 3.1). The finite size of the observable volume results from the fact that our cosmos has a finite age (of 13.7 billion years) and that light and other information can only reach us from distances for which the light travel time has been smaller than the age of the universe. As a matter of fact, the distance from which we can receive light is even smaller, since during the first 380 000 years Thomson scattering by the free electrons of the hot cosmic plasma made the cosmos opaque. Only when the electrons in the expanding and cooling plasma recombined with the atomic nuclei, the radiation field decoupled from the matter and the cosmos became transparent for most wavelengths. Obviously, only light emitted after this epoch (i.e., only light which was emitted after the cosmos had reached an age of about 380 000 years) can reach us. This includes the photons which were emitted by the recombining plasma itself. At the epoch when the cosmos became transparent, the plasma had a

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_1, © Springer-Verlag Berlin Heidelberg 2009

temperature of about 3×10^3 K. Thus, mainly visual and near-infrared photons were emitted by the plasma. But due to their redshift of $z \approx 10^3$ these photons today are observed as the Cosmic Microwave Background (CMB), which was discovered in 1965 [408] at radio wavelengths.

Since the plasma emitting the CMB is the most distant source from which light can reach us, in the light of the CMB we directly overlook the whole cosmic volume which in principle can be observed with astronomical techniques. For all other types of astronomical objects the volume which we can overlook is smaller. There are two reasons for this: Firstly, all other known astronomical objects were formed at later epochs. Thus, their light was emitted later and at a smaller distance from us. Secondly, for many types of objects the most distant existing examples are too faint to be detectable with present technologies. Only a few types of intrinsically very bright sources can be observed at practically all distances where they are expected to exist. Among these intrinsically very bright sources, where we have a complete overview of the total volume where they theoretically can be observed, are the most luminous quasars. Using absorption features which they produce in the spectra of luminous quasars, intergalactic gas clouds and the gaseous disks of gas-rich galaxies can be studied in practically the same volume. Our possibilities of observing the most distant galaxies are somewhat more restricted. As pointed out in the following chapters, we are very close to observing luminous galaxies up to the distance corresponding to the epoch where the first of these objects have been formed. However, galaxies (probably) formed earlier than the luminous quasars, and (as will be explained later) the first (and thus most distant) galaxies were relatively small and faint systems. Therefore, their detection is technically more demanding than observing distant quasars.

In light of the CMB we see a universe consisting of an almost homogeneous hydrogen–helium plasma without heavy chemical elements and with a relative density variation of the visible matter of only about 10^{-5}. Obviously, at this epoch the universe was highly homogeneous on all scales. It was very different from today's structured universe, which contains galaxies, stars, planets, and a multitude of chemical elements, including the heavy elements which are the main constituents of our planet, and of ourselves. This structured and complex universe emerged several hundred million years after the emission of the microwave background when the first stars and galaxies started to form due to the gravitational contraction of the tiny density fluctuations visible in the CMB. The technical progress of the last two decades has made it possible to directly observe this interesting cosmic transition period and to derive, at least in principle, information on the very early galaxies. Obviously, the properties of the early galaxies contain important information of the initial conditions for the formation of the present universe. Moreover, this data is expected to provide constraints on the physical processes in the early universe.

In principle, information on these initial conditions can be derived also from observations of the local universe, since all the properties and the structure observed in our present cosmos have evolved under the laws of physics from these early stages. However, at present there are still large gaps in our knowledge of important constituents of the cosmos (such as the dark energy and the dark matter), and

many of the relevant astrophysical processes (such as star formation and galaxy interactions) are not well understood. This makes conclusions based on a backward extrapolation from observations of the present-day evolved universe inaccurate and unreliable. Direct observations of the high-redshift universe are always preferable. Moreover, where extrapolations are unavoidable, observations extending to high redshifts provide a more reliable basis for such extrapolations into the unseen past.

1.2 Space, Time, and Redshift

As pointed out above, observing distant astronomical objects always means observing the past of the universe. The objective of such studies is to obtain information on the history and evolution of our world.

Since the cosmos is not completely homogeneous, there are small local variations of the space curvature, and clocks actually run differently at different locations. However, if the cosmic density is averaged over large scales, the relative density variations become progressively smaller with increasing scale. Thus, on large scales the universe appears homogeneous. Therefore, it is possible to define a mean cosmic time, which (except for regions of very strong gravitational fields) is a good approximation for all reference frames which are locally at rest.

From a variety of astronomical observations we know that our universe is expanding. (A listing and a critical discussion of the evidence for the cosmic expansion can be found, e.g., in [446].) The cosmic expansion means that the distance between objects, which are locally at rest, is increasing steadily. It also means that the wavelength of light is increasing while traveling between two objects which are locally at rest. The ratio between the wavelength of the observed light and the wavelength of the emitted light is equal to the relative change of the linear distance scale, i.e.:

$$\frac{\lambda_{obs}}{\lambda_{em}} = \frac{a_{obs}}{a_{em}} \tag{1.1}$$

where λ_{obs} is the observed wavelength, λ_{em} is the emitted wavelength, and a_{obs}/a_{em} is the ratio between the cosmic scale when the light is observed, and the cosmic scale at the time of the light emission.

In astronomy this wavelength change is usually expressed by the "redshift" z, which is defined as

$$z = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} = \frac{a_{obs}}{a_{em}} - 1 \tag{1.2}$$

For spectral lines, λ_{em} can be determined by means of laboratory measurements or by quantum-mechanical calculations.

In the older astronomical literature, redshifts are often expressed by the product cz (where c is the velocity of light) and are listed in units of km s^{-1} (often without explicitly specifying the units).

The wavelength change actually measured in astronomical spectra differs from the redshift as defined above, since, due to their mutual gravitational attraction and variations of the local mass density and gravitational potential, galaxies normally are moving relative to the local rest frame. Due to the Doppler effect, these local motions (called "peculiar velocities") also produce wavelength shifts, which add to the wavelength shifts due to the cosmic expansion. In order to derive accurate redshifts caused by the cosmic expansion, the measured wavelength shifts have to be corrected for the peculiar velocities of the observed objects (and for the motion of the observer, including the rotation of the Earth, the orbital motion of the Earth, the motion of the solar system in the Galaxy, and the peculiar velocity of the Milky Way galaxy). The effects of the peculiar velocities normally amount to $\Delta\lambda/\lambda \leq 0.01$. Therefore, the relative effect of the peculiar motions are insignificant for large redshifts, but corrections for peculiar motions are very important for nearby objects.

Since the redshift as defined above increases monotonically with distance and with the time which elapsed since the observed light was emitted, in the astronomical literature the dimensionless redshift z is often used as a measure of distances, for characterizing ages, and for defining time intervals. The use of z obviously has the advantage that, at least in principle, z is a directly measurable quantity, and that distances and times expressed in terms of z can be easily compared. However, in practice the correction for peculiar velocities is not always easy or reliable. Moreover, in order to convert z into a linear distance or time units (such as meters, parsecs, seconds, or years)[1] the correct world model and the cosmic parameters have to be known (see Sect. 3.1).

1.3 Historical Remarks

Astronomical observations were among the first scientific activities in the history of mankind. From archaeological evidence we know that systematic observations of the Sun and of the night sky were carried out already during the Stone Age. These early astronomical observations were probably always part of religious ceremonies and divinations. However, their success required the the development of basic scientific methods.

Many prehistoric archeological sites show an exact alignment with respect to astronomical objects or events, indicating that accurate position measurements of celestial objects were achieved already at very early epochs. On the other hand, the development of methods to derive astronomical distances took much longer. Such

[1] The most commonly used unit for measuring astronomical distances is the "parsec" (pc). It is defined as the distance from the Sun at which the semimajor axis of the Earth's orbit has an angular extent of one arcsecond (arcsec). In SI units we have 1 pc $\approx 3.086 \times 10^{16}$ m. Large distances are expressed in kpc, Mpc, and Gpc.

measurements became possible only after the Ancient Greeks found that distances can be determined by measuring angles in a triangle in which the length of one side (the baseline) is known. Using this method, the distance of an object can be derived by measuring the angular difference (called parallax) between the two directions under which the object is observed at the two end points of a known baseline. Since initially all astronomical distances were based on the measurement of such angles, in astronomy the term parallax has become synonymous for the distance itself. Thus, in the astronomical literature, a "parallax" can be an angle (measured in radians or arcsec) or a distance (measured in km or pc).

Among the baselines used for astronomical parallax measurements are known distances on the Earth, the Earth radius, the radius of the Moon's orbit, and the diameter of the Earth's orbit around the Sun. An approximate value for the Earth radius was determined already by Eratosthenes (276–194 BC). Using the Earth radius as a baseline Ptolemy succeeded around AD 150 in measuring the parallax and the distance of the Moon with a relative error $<2\%$. However, the Ancient Greek's attempts to derive the distance of the Sun, using the Moon's orbit as a baseline, resulted in a much too small solar distance. Since the main error source (atmospheric refraction) was not understood until much later, even at the time of Kepler the solar distance and the size of the solar system were still not well known. Only later in the 17th century the correct order of magnitude of the distance to the Sun could be established by Wendelin, who in 1650 derived a value which was correct within a factor of two.

It took almost another two centuries before Bessel in 1837, using the Earth's orbit as a baseline, succeeded in deriving the parallax of a star (61 Cyg). This first distance to an object outside the solar system and accurate lower limits for distances to other galactic stars also provided a first indication of the size of the Milky Way galaxy.

That the visible universe extends beyond the Milky Way galaxy, was not yet clear at that time. Since the invention of the telescope, spiral galaxies had been observed and investigated. That these diffuse objects were outside our Milky Way system and were large stellar systems of their own, had been suggested and discussed since their discovery. But first attempts to measure their distances lead to contradictory results. A reliable confirmation of their extragalactic nature became possible only during the first quarter of the 20th century, when objects of known absolute luminosity (such as bright stars, variable stars, and novae) were observed in the Andromeda nebula and in other Local Group galaxies [117, 239, 240, 322, 389, 473] establishing their large distances.

Among the methods which had been used to derive the distance of the Andromeda nebula was the period-luminosity relation for pulsating stars (see, e.g., [447]). This relation had been found empirically by H. S. Leavitt in 1908 [304, 419] for the δ Cephei variable stars (also called "Cepheids"). Edwin Hubble [239] detected variable stars with the characteristic light curves of Cepheids in the Andromeda nebula. Comparing their observed apparent brightness with the absolute luminosity derived from their pulsation periods and using the period-luminosity relation accepted at

that time, he obtained their distances and, thus, the distance of their host galaxy. Soon similar results were obtained for several other nearby extragalactic systems. But with the photographic plates used at that time, it was not possible to observe δ Cephei variables in galaxies that were significantly more distant than Andromeda.

log Dm	Argum. log Dm v	n
0.24	+827 km	9
0.43	+656	7
0.66	+512	8
0.88	+555	10
1.07	+334	5
1.71	— 20	3

Fig. 1.1 Table published by C. Wirtz in 1924 showing the existence of a relation between the mean angular diameter Dm (in arcsec) and the mean redshift of galaxies. The redshifts are expressed as apparent Doppler velocities $v = c\,\Delta\lambda/\lambda = c\,z$ and given in km s^{-1}. Obviously z increases with decreasing mean angular diameter Dm. Wirtz based his table on 42 galaxies, grouped into 6 bins. n is the number of objects in each bin. He correctly concluded that the observed correlation results from an increase of the redshift with distance. However, since no reliable extragalactic distance measures were available in 1924, Wirtz was not able to calibrate his qualitative redshift-distance relation. (Reproduced from [562])

Several years before the physical nature of the spiral nebulae had been clarified, the systematic redshift of their spectra had been found by V. M. Slipher [480]. Using (mainly) Slipher's data C. Wirtz published in 1922 and 1924 two important papers in which he showed that these redshifts were correlated with the apparent total brightness and with the angular diameters of the observed galaxies (Fig. 1.1). Arguing correctly that, on average, the apparent brightness and the angular diameter of galaxies are expected to decrease with distance, he concluded that the observed correlations were due to the existence of a redshift-distance relation for galaxies [561, 562]. Although Wirtz was probably the first one to notice the distance dependence of the galaxy redshifts, he was not able to calibrate this relation, as his distance indicators gave relative distances only. It took another 5 years until Edwin Hubble in his fundamental 1929 paper [241] was able to present a calibrated redshift-distance relation, which today is referred to as the "Hubble diagram" (Fig. 1.2). To derive this diagram Hubble used a sample of about 30 galaxies with known redshifts. For the nearer galaxies in his sample he determined the distance using δ Cephei variables. For the more distant galaxies Hubble used various indirect calibrators, assuming that the distant galaxies had intrinsic properties similar to those found for the nearby galaxies in which Cepheids had been observed.

Since Hubble's (1929) distance estimates contained large systematic errors, the numerical value of the slope of his redshift-distance relation (the "Hubble constant")

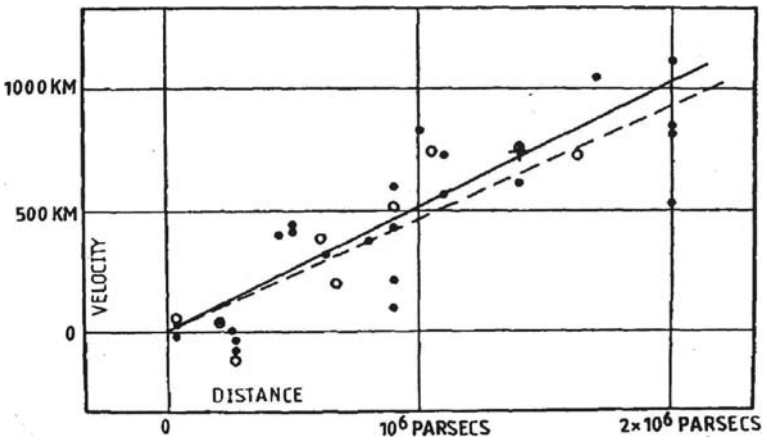

Fig. 1.2 The original Hubble diagram of 1929. The plot gives the observed redshift (again expressed as a Doppler shift in km s^{-1}) as a function of the distance (in parsec). (Reproduced from [241])

was about seven times too large. Consequently, distances derived from the initial Hubble relation were about seven times too small. Thus, the first extragalactic distances were about as inaccurate as the first solar distance derivations by the Ancient Greek astronomers. It took about 30 years to bring the Hubble constant to within a factor two of its correct value, and it took the rest of the 20th century to reduce the relative error to about 10% . (For details of the interesting history of the Hubble constant see, e.g., [508, 525].)

As seen from today, Wirtz and Hubble discovered the expansion of the universe and provided the foundations of present-day cosmological theories. However, neither Wirtz nor Hubble was aware of this discovery. Both regarded their results as support for W. de Sitter's [128] cosmological model, which assumed a non-expanding, static world, where a galaxy redshift results from a distance-dependent time. Although expanding world models based on Einstein's theory of general relativity had been developed by A. A. Friedmann [181, 182] already in 1922, it took much time before the astronomical community took note of these models. Like Einstein and de Sitter almost everybody was still convinced that the universe was static. Even after additional measurements had shown that the observed redshift-distance relation was quantitatively incompatible with the prediction of de Sitter's model, the expanding Friedmann-Lemaître Big-Bang world models found little acceptance in the scientific community. Other cosmological models, such as "tired light" theories and the Steady State cosmology were initially much more popular. Only during the second half of 20th century it became almost generally accepted that we live in a universe which is expanding according to Friedmann's solutions of Einstein's equations.

Although its initial calibration was rather inaccurate and although its physical basis was unclear, Hubble's relation immediately became the most important tool

for estimating extragalactic distances. But many astronomers were well aware of the uncertainties of the calibration of the redshift-distance relation. Moreover, the cosmic redshift-distance relation is linear only for small distances. For larger redshifts a conversion of the measured redshifts into distances requires the knowledge of the expansion law and the cosmic parameters. But independent of the calibration problem, redshifts provide at least a reliable relative measure of large distances.

The galaxies in Hubble's 1929 sample all had redshifts $z < 0.004$. But already during the year of Hubble's important publication Humason and Pease observed a galaxy redshift as large as $z = 0.026$ [243]. Two years later the record stood at $z \approx 0.066$. In 1949, after the Mt. Palomar 200-inch telescope had become operational, the high-redshift frontier was pushed to $z \approx 0.203$. During the following years progress slowed down since the low surface brightness of galaxies with higher redshifts (see Sect. 3.1.5) made it difficult to obtain spectra of such objects on photographic plates. By 1960 the radio galaxy 3C295 ($z = 0.462$) observed by Minkowski [355] became the most distant object known. During the following four decades radio sources (radio galaxies and quasars) remained the objects with the largest observed redshifts. Although the first quasars, which were discovered in 1963, had moderate redshifts, within the following 2 years quasars with $z > 2.0$ became known. By 1982 many quasars with redshifts >3 had been discovered and the record stood at $z = 3.87$. In 1999 quasars redshifts of 5.0 were reached [166], and today (as a result of the SDSS survey) a significant sample of quasars with redshifts $z > 6$ is known (e.g., [167, 168]).

The first normal galaxy with a very high redshift ($z = 3.215$) was discovered by Djorgovski et al. [140] while looking for Lyα emission objects near the quasar PKS1614+051. During the following years similar detections were reported by Schneider et al. [459], Hu and Cowie [235], and Steidel et al. [499]. These observations were important as they demonstrated that there existed galaxies at these redshifts which were bright enough to be observed and studied with the techniques available. As a result of these early observations, systematic searches were started which, together with a new class of telescopes with apertures of 8–10 m, yielded within a few years large samples of well-identified galaxies with redshifts $2 < z < 6$, thereby opening the field of this book. During the past years galaxies with redshifts of at least $z = 7.6$ (and possibly >9) have been identified (e.g., [68]), while the highest observed QSO redshift stands at $z = 6.4$.

Chapter 2
The Nearby Universe

Due to their large distances, high-redshift galaxies are characterized by small angular diameters (typically not more than a few arcsec), very weak light fluxes, and a low surface brightness. Therefore, when observing such objects, one normally has to make do with low spatial and spectral resolutions, and (due to photon noise) with low signal-to-noise ratios. As a result, the direct information content of such data is always rather limited. In order to evaluate these observations, they are often compared to observations of galaxies with known physical properties in the local (or present-day) universe. The local galaxies can be studied in much more detail and with higher accuracy, and their physical properties can be determined directly. The properties of the distant galaxies are then often inferred from comparing their observations with those of local galaxies, or by fitting the observations to models which are based on data obtained in of the local universe. This approach assumes that the basic physical processes in distant galaxies are sufficiently similar to those in the local universe to make such comparisons meaningful. Since reliable tests of this assumption are difficult, the experts working in the field sometimes passionately disagree on whether this assumption is justified and on whether properties of local galaxies can be adopted for interpreting high-redshift data. But, in spite of sometimes conflicting opinions, most recent results on high-redshift galaxies have been based on the above assumption and on comparisons with galaxies in the nearby universe. Therefore, before discussing the observations of high-redshift objects, in the present chapter our knowledge and our assumptions concerning the present-day or "local" universe will be summarized.

A close look at the local environment is also useful since some of the processes and events which took place in the early universe left long-lived remnants and traces which can be studied locally. Some of these remnants contain important information on the early universe. An example for such "fossils" of the early universe are old stars of very low mass, which evolve on a timescale longer than the age of the universe.

A final reason for looking at the local universe in the context of high-redshift objects is the fact that any correct cosmic theory must result in the universe as we see it today. Thus the local universe forms an important boundary condition and constraint for cosmological models.

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_2,

The following brief description of the local universe starts with our immediate neighborhood and ends with the largest distinct objects and structures which are known to us at present.

2.1 Planets, Stars, and Black Holes

From the standpoint of astronomy our cosmic home is the surface of a small planet, which is orbiting a low-mass star (our Sun), which in turn belongs to a large spiral galaxy (the Milky Way galaxy).

2.1.1 Planets

In modern astronomy planets are defined as spherical, or nearly spherical, hydrostatic objects of substellar mass, which are orbiting a star. Until about 1995 only the planets of the solar system were reliably known. Since then searches for extrasolar planetary system resulted in the discovery of many planets around other stars. The high success rate of these searches indicates that planetary systems are a rather common property of stars.

In addition to planets, planetary systems can also include smaller bodies (such as dwarf planets and comets) and various types of interplanetary matter. But (at least in the solar system) the small bodies do not contribute much to the total mass. More information on the solar system and the exact current definition of its components can be found on the web page of the International Astronomical Union (www.iau.org).

From our solar system we know that individual planets can have very different characteristics. But in all known planetary systems the total mass of all planets amounts to only a tiny fraction of the mass of the central star. Thus, although planets are common, their contribution to the total cosmic mass is negligible. On the other hand, to the best of our present knowledge, planets and their satellites are the only locations in the universe where life (of the form we know) appears to be possible. Therefore, planets in the nearby universe are obviously of high interest. Almost certainly, planets are also orbiting the stars of high-redshift galaxies. But the detection of these distant planets is far outside our present capabilities. And, in view of the time it took to develop life on Earth (see, e.g., [532]), the light from the high-redshift galaxies discussed in this book probably was emitted long before life had a chance to develop there.

2.1.2 Stars

2.1.2.1 Basic Properties

Scientifically, stars are defined as spherical or nearly spherical objects in hydrostatic equilibrium, having (in the local universe) masses M approximately in the range

$10^{-2} < M/M_\odot < 10^2$, where $M_\odot = 1.989 \times 10^{30}$ kg is the mass of the Sun. Hydrostatic equilibrium means that anywhere in a star the gravitational forces pulling the matter toward the stellar center are just compensated by the forces resulting from the local pressure gradient.

Stars with $M > 100\,M_\odot$ are extremely rare in the local universe, and stars more massive than $150 M_\odot$ have not been found so far. The exact lower stellar mass limit is less well known at present. But stars of very low mass are not expected to contribute much to the light or the mass of the observable high-redshift galaxies. Therefore, stars of very low mass will not be discussed here.

Stars are formed by the gravitational contraction of dense, cool interstellar clouds (see Sect. 2.2.5). In all galaxies containing cool interstellar gas, star formation still takes place. In our Milky Way galaxy, at present, interstellar gas is converted into stars at a rate of about one $M_\odot$ per year. As discussed in Sect. 2.2, the star formation rate (SFR) differs greatly between different galaxies. In some galaxies star formation has stopped already a long time ago. Other galaxies have (temporarily) very large SFRs, sometimes exceeding $10^3\,M_\odot\ \mathrm{yr}^{-1}$.

In the astronomical literature, individual stars are normally referred to by a catalog number or by their coordinates on the sky. Most frequently the catalog numbers of the Henry-Draper catalog (HD) are used. A few bright stars have individual names. Catalog numbers and stellar names can be converted into coordinates using the Internet facilities of the large astronomical data bases, such as the CDS (http://cdsweb.u-strasbg.fr).

2.1.2.2 The Distribution of the Stellar Masses

Except near the upper and lower mass limits, the distribution of stellar masses of young stars in our galaxy can normally be described by the "Salpeter initial mass function"

$$\frac{\mathrm{d}N}{\mathrm{d}\log M} \propto M^{-\chi} \tag{2.1}$$

where $\chi \approx 1.3$ for $M > 0.5 M_\odot$ and $\chi \approx 0$ for $M < 0.5 M_\odot$. Such a law has first been suggested by E. E. Salpeter ([445]). A discussion of the observational constraints for the parameters of the Salpeter mass function can be found, e.g., in [285] and [286]. The Salpeter initial mass function and related expressions are most widely used to characterize the stellar mass distribution of newly formed stellar populations.

Since stars loose mass during their evolution, and since stars of different mass have very different life expectancies, the mass distribution strongly evolves with the age of a stellar population. As a result the stellar mass distribution of a galaxy depends on the initial mass function as well as on the star formation history of the system. The (present) mass distribution of stars in the solar neighborhood is described, e.g., in [92].

During their evolution, all stars lose mass due to outflows from their surface layers. These outflows (called "stellar winds") can be driven by thermal pressure (as

in our Sun) or by radiation pressure in resonance lines (as in the case of hot stars) or by radiation pressure on dust grains (in very cool stars with low surface gravities). The individual mass loss rates depend on the mass, on the evolutionary phase, and on the chemical composition of the stars. Particularly high-mass loss rates are found for very massive stars (where values up to 10^{-4} $M_\odot$ yr^{-1} are observed) and in late phases of the evolution of low-mass stars.

2.1.2.3 Chemical Composition

Since stars form from interstellar matter, their initial chemical composition is that of the local interstellar gas. Thus, all stars initially contain mainly hydrogen (mass fraction $\approx 70\%$) and helium ($\leq 30\%$). Heavier elements contribute less than 3% of the initial stellar mass. During the stellar evolution, nuclear reactions in the interior and the loss of hydrogen-rich surface layers by stellar winds result in a gradual decrease of the hydrogen fraction. In some stars this leads to a complete loss of all the hydrogen. Important examples of such hydrogen-free stars are the massive and hot Wolf-Rayet or "WR" stars (e.g., [115]) which in their spectra show strong spectral lines of helium, carbon, nitrogen, and similar element but no hydrogen lines (Fig. 2.1). Due to their distinct spectra and their high luminosities, WR stars can be detected even in the most distant galaxies.

2.1.2.4 Stellar Spectra

Since stars are very opaque, we receive light only from their surface layers. The layer where the observed photons are emitted is called the "stellar photosphere." In stellar spectra we observe the light emission and absorption of the photospheric gas. Since the stars are optically thick, the stellar continua show on first approximation a Planck spectrum. On this continuum spectral lines are superposed. Normally the temperature of the stellar gas is decreasing toward the surface. Therefore, the radiation intensity normally is higher in the continuum (where we can look deeper into the star) than in the spectral lines, which (because of the higher line absorption cross section) originate from higher and cooler stellar layers. Thus, stellar spectra characteristically show spectral lines in absorption. (In our Sun these absorption lines are known as "Fraunhofer lines".) However, there are exceptions. Many stars (including our Sun) have above the photosphere a temperature minimum, followed by an increase of the temperature in the optically thin atmospheric layers above the photosphere. The (weak) lines originating in these hot layers occur as emission lines. Emission lines are also produced in extended, optically thin atmospheres, in stellar winds, and in circumstellar gas masses.

Both the continua and the line spectra of stars depend on their (effective) surface temperature, on the gas density at the surface, and on the chemical composition. For non-degenerate stars with a normal chemical composition, the most important parameter for the spectral appearance is the effective temperature T_{eff}. The observed effective temperatures cover approximately the range between 2000 K and 10^5 K. Depending on T_{eff}, the photospheric spectra are dominated by absorption lines of

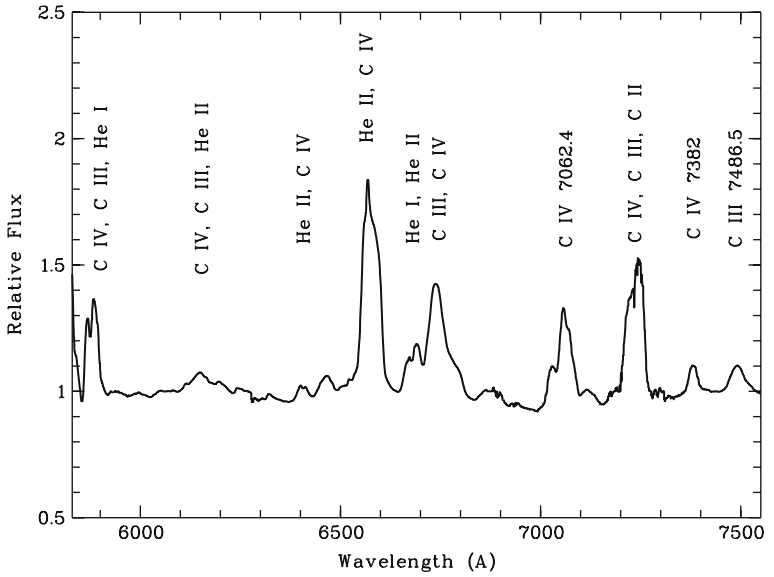

Fig. 2.1 Spectrum of the Wolf-Rayet star γ^2 Velorum. Due to strong stellar winds this star has lost all of its original hydrogen-rich surface layers. Therefore, matter from the stellar interior is now observed at the stellar surface. Since in the interior all hydrogen has been destroyed by nuclear reactions, no hydrogen lines are visible. Instead the spectrum is dominated by spectral lines of elements (such as helium and carbon) which have been synthesized in the nuclear processes. These lines appear as broad emission blends since they are formed in the still active strong stellar wind of this hot and luminous star. (Spectrum courtesy O. Stahl)

molecules, of neutral atoms, or of ions of different ionization stages. At very low temperature dust emission can play an important role.

In very luminous stars (which are particularly important for the appearance of the spectra of high-redshift galaxies) the observed spectra are often a superposition of weak photospheric absorption lines, the spectral signatures of the stellar winds, and of sharp interstellar absorption lines, which are produced by the interstellar gas along the line-of-sight to the star. This is illustrated by Fig. 2.2, where the UV spectra of several hot high-luminosity stars are plotted.

2.1.2.5 Spectral Types

In the astronomical literature the stellar spectra are often characterized using "spectral classes" or "spectral types." In the context of the stellar populations of galaxies, quantitative parameters (such as the stellar luminosity, the effective temperature, the stellar radius, and the chemical composition) are more useful than spectral classes. However, the spectral types provide a convenient tool to characterize stellar spectra in a quick way, and they are widely used in the astronomical literature. Therefore, the most basic properties of the Morgan-Keenan-Kellman (MKK) or Morgan-Keenan (MK) classification scheme [363] have to be included here.

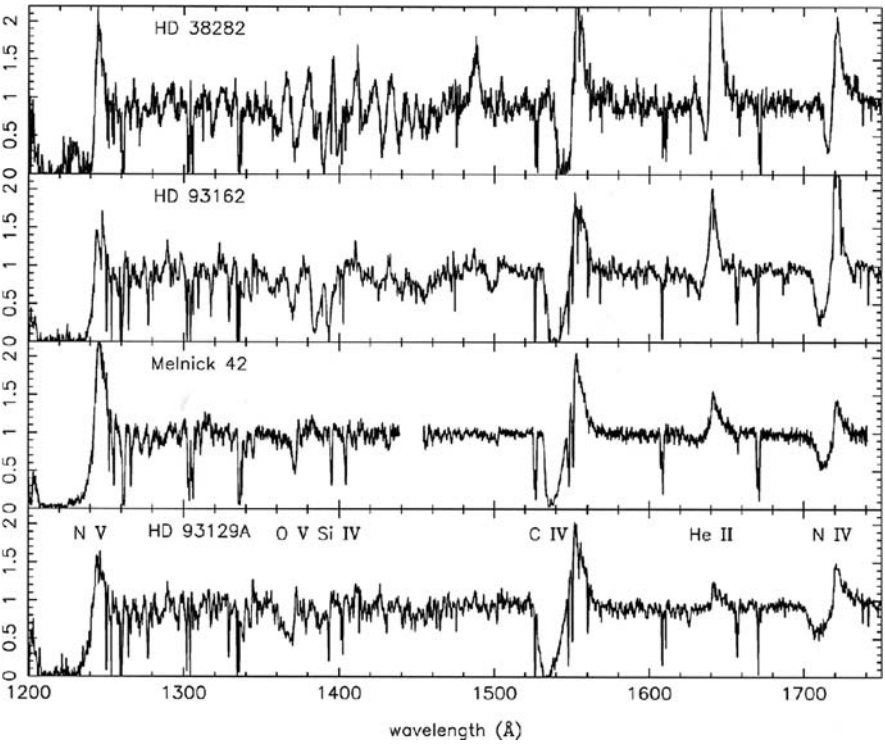

Fig. 2.2 UV spectra of four very luminous and hot stars. The star HD 93129A (*bottom*) is an O2If star with an (essentially) normal galactic chemical composition. (In the classification of O stars the "f" indicates the presence of emission lines.) The stars HD 38282 (*top*) and HD 93162 are Wolf-Rayet stars (of types WN6-A(B) and WN6-A, respectively). The star Melnik 42 has a type between WN and O2. The spectral features identified in the figure are the stellar-wind lines of N V 1239 Å, 1243 Å, O V 1341 Å, Si IV 1394 Å, 1403 Å, C IV 1548 Å, 1551 Å, He II 1640 Å, and N IV 1718 Å. Note the very broad, blueshifted stellar-wind absorption features in the spectrum of the luminous O-type star. The sharp lines in the spectra are of interstellar origin. (From Walborn et al. [549])

In the MK classification system the stars are ordered in an effective-temperature sequence using the letters O, B, A, F, G, K, and M, where the letter O represents the hottest stars and the letter M the coolest stars. In the modern literature, stars cooler than M (as defined in the MK system) are classified as L and T. The individual temperature classes are subdivided by numbers between 0 and 9. Moreover, the star's luminosity is indicated by a Roman number between I (for a very high luminosity) and V (for hydrogen-burning "main-sequence" stars). The full MK spectral type of our Sun is G2V. Not all subclass numbers are actually used. There are, e.g., no stars of class O0 and O1 defined. The hottest stars with a normal chemical composition have the spectral type O2.

For stars with atmospheres of unusual chemical composition (such as the Wolf-Rayet stars mentioned above), special classification schemes have been defined (see,

e.g., [115]). Moreover, extensions and refinements of the MK system have been defined by W. W. Morgan and his colleagues (and later by other authors). A discussion of these refinements is outside the scope of this book. But as a word of warning, it must be noted that due to the updates and refinements of the MK system, for a given star in the literature different classifications may be listed.

2.1.2.6 Modeling Stellar Structure and Evolution

In order to solve the equations describing the hydrostatic equilibrium of a star, one has to know the equation of state (i.e., the relation between the pressure and the density) and other material properties of the stellar matter. In most stars the interior pressure follows with good approximation the pressure law of an ideal gas (i.e., the pressure is proportional to the density and the temperature). In hot, massive stars the pressure of the radiation field often dominates. In this case the pressure depends on the temperature only, being proportional to the fourth power of the temperature. In the dense cores of certain cool stars and in very compact stars (such as white dwarfs and neutron stars), the Pauli principle results in deviations from the Boltzmann momentum distributions of the fermions. With increasing density the fermions have to populate progressively higher momentum levels. This results in a progressively higher pressure, which depends essentially only on the density (and becomes independent of the temperature). The pressure of such a "degenerate" fermion gas can be much higher than in an ideal gas of the same temperature. In the case of strong degeneracy, the effect of the gas temperature becomes negligible. For a completely degenerate non-relativistic electron gas, the pressure becomes proportional to $\rho^{5/3}$, where ρ is the density.

In a normal high-density stellar plasma the free electrons (as the lightest fermions) are most susceptible to degeneracy. Therefore, in most high-density stars the pressure is essentially determined by the degenerate electron gas. An exception are the neutron stars, which have interior densities $> 10^{17}$ kg m^{-3}. At these densities the electron energies reach values which are sufficiently high to result in the formation of neutrons by the recombination of electrons and protons. As a result, the electrons are removed and degenerate neutrons (and hyperons) dominate the pressure.

Except for the neutron stars, the equations of state of stellar matter are well understood. For an ideal gas, a plasma dominated by radiation pressure, a degenerate electron gas, and any combination of these components, the pressure can be readily computed, if the density and temperature are known.

Stars have simple geometric forms (spherical or nearly spherical) and (in contrast, e.g., to planets) the equation of state of stellar matter is relatively well known. Therefore, the structure and the time evolution of stars can be modeled rather reliably by numerically solving the equations describing the conservation of mass, momentum, and energy, taking into account the energy transport, and starting from plausible initial conditions. A detailed overview of the stellar structure theory and the basic numerical methods for modeling stellar structure and evolution can be found, e.g., in [268]. The special case of degenerate stars is discussed, e.g., in [86] and [223].

2.1.2.7 The Hertzsprung–Russell Diagram

In the non-degenerate stars the high temperature of the gas (or of the radiation field) required to maintain hydrostatic equilibrium normally is due to the energy input of nuclear reactions taking place in the hot interior. In most stars (including our Sun) the dominant reaction is the conversion of hydrogen into helium (referred to as "hydrogen burning" in the astronomical literature). The predominance of the H → He conversion results from the high energy yield of this reaction and from the fact that hydrogen is the most abundant chemical element. All stars massive enough to ignite hydrogen in their center spend most of their lifetime in the hydrogen-burning phase. In the so-called "Hertzsprung–Russell diagram" (or "HR diagram"), where stellar luminosities L are plotted as a function of the (effective) surface temperatures, the hydrogen-burning stars form a conspicuous linear sequence (Figs. 2.3 and 2.4). Therefore, stars burning hydrogen in their cores are called "main-sequence stars." A main sequence is also visible in a related diagram, where the surface gravitational acceleration is plotted as a function of the surface temperature of stars (Fig. 2.4, lower panel). Main-sequence stars are also referred to as "dwarf stars," while stars located in the HR diagram above the main sequence are called "giants" or "supergiants." The exact position and form of the main sequence in the HR diagram depends on the chemical composition and on the age of the stars. Therefore, the main sequence is particularly well defined for members of star clusters which consist of stars of the same age and the same chemical composition. For comparison purposes, often a hypothetical "zero-age main sequence" (ZAMS) is calculated, where all stars have just started hydrogen burning. Since the pre-main sequence evolutionary timescale of stars depends on the mass, stars which form together in a cluster, reach the onset of hydrogen burning at different times. Thus, an observed main-sequence always differs from the theoretical ZAMS.

The HR diagram forms an important interface between the observations of stars and the theory of stellar structure. If the distance is known, stellar luminosities can be directly observed. Effective surface temperatures can be determined from stellar spectra. For faint stars, where spectra of sufficient quality cannot be obtained, approximate effective temperatures can be deduced from flux ratios at different wavelengths. In astronomy these flux ratios (usually expressed in magnitudes, see Sect. 4.2.4) are called "color indices" or "colors." These colors can be calibrated in terms of effective temperatures by means of stellar models (as in Fig. 2.3). While the calibration is well understood and reliable for most main sequence stars, there still exist uncertainties for the cool stars of very low mass.

For a given initial chemical composition the parameter varying along the main sequence in an HR diagram is the stellar mass. Both the temperature and the luminosity of the main-sequence stars increase monotonically with the stellar mass. As a consequence, there exists a mass-luminosity relation for core hydrogen-burning stars. This relation is normally expressed as

$$L = C M^{\eta} \tag{2.2}$$

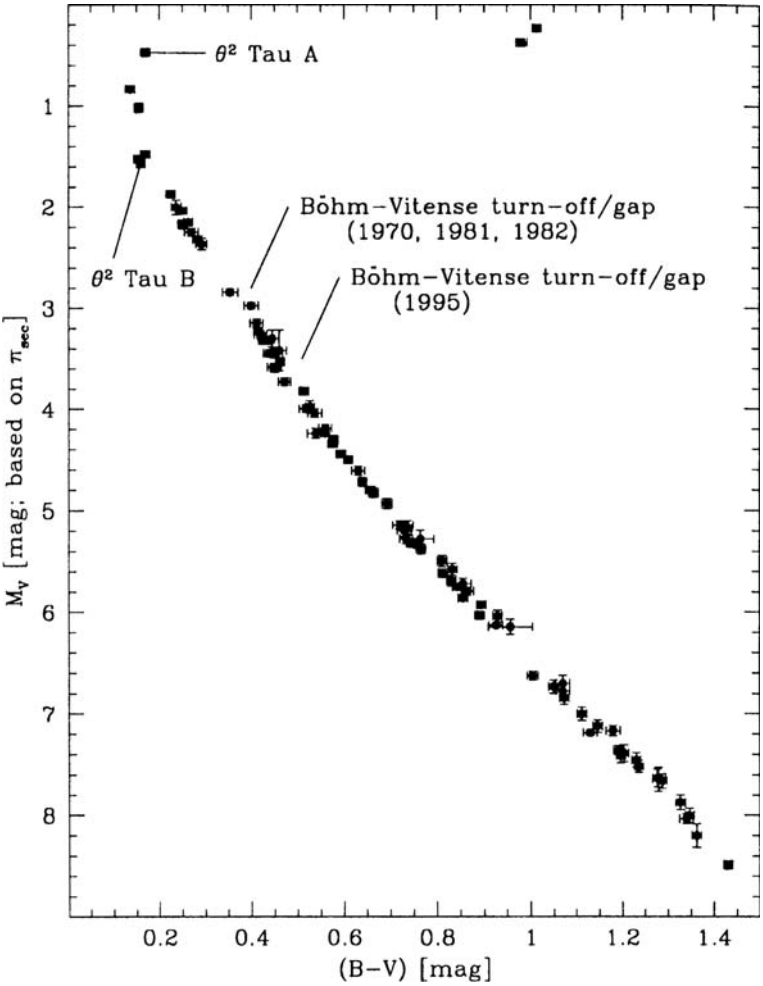

Fig. 2.3 HR diagram of the members of the Hyades star cluster. The effective temperature is expressed by the color index $B - V$ (see Sect. 4.2.4). The luminosity is given in magnitudes. All luminosities are based on very accurate distances derived geometrically with the Hypparcos astrometric satellite. Note the very well defined "main sequence" of the core hydrogen-burning stars, which (being cluster members) all have the same age. The two stars far above the main sequence are "giants" which have already reached the core helium burning evolutionary stage. The gaps in the main sequence labeled "Böhm–Vitense turn-off/gap" are due to qualitative changes of the surface physical conditions of main-sequence stars at the corresponding effective temperatures. (From de Bruijne et al. [126], where further details are described)

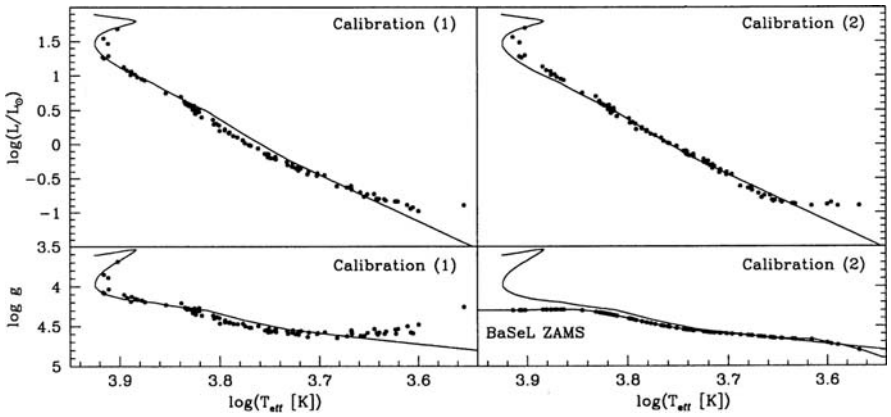

Fig. 2.4 HR diagrams and $g_{eff}(T_{eff})$ diagrams of the same stars as in Fig. 2.3, but using the effective temperature as abscissa. The *right* and the *left panels* are based on two different calibrations using different sets of stellar models. The luminosity is expressed in units of the solar luminosity $L_{\odot}$, the temperature is given in K. The effective gravitational acceleration g_{eff} is given in cm s^{-2} The *solid lines* are computed main sequences based on stellar models. The deviations at very low effective temperatures indicate systematic errors in the model computations at these temperatures. (From de Bruijne et al. [126])

where C is a constant (which depends on the initial chemical composition), and where (on average) $\eta \approx 3.5$ [268]. Thus, the luminosity of hydrogen-burning stars increases strongly with mass. On the other hand, the amount of nuclear fuel available for a main sequence star is roughly proportional to its mass.[1] Consequently, the duration of the hydrogen-burning phase decreases with the mass roughly proportional to $M/L \sim M^{-2.5}$. While our Sun spends more than 10^{10} years in the main sequence stage, for the most massive stars core hydrogen burning lasts about 10^6 years only. Stars which are significantly less massive than the Sun have main-sequence lifetimes in excess of the present age of the universe. Thus, post-main-sequence stars of such low mass do not exist.

Main-sequence stars with masses above about $60M_{\odot}$ are vibrationally unstable [305, 462]. However, due to shock damping the amplitudes of the pulsations resulting from this instability is always small [14, 15]. Thus, the vibrational instability cannot cause the upper mass limit of stars observed in the local universe. A more likely explanation for the absence of stars with masses above about $150M_{\odot}$ is that (under the conditions of the present universe) the formation of such objects is prevented by the ignition of hydrogen burning in the protostellar cores (e.g., [20]). However, because of the complexity of the star formation process, the present theory of high-mass star formation (e.g., [579]) does not yet allow us to predict the stellar upper mass limit reliably.

[1] Since the stars loose part their initial mass by stellar winds, and since the fraction of the hydrogen mass that can be burned depends on the initial mass, the relation between stellar mass and the mass of the available fuel actually is more complex.

2.1.3 The Evolution of Low-Mass Stars

Stellar evolution depends critically on the initial stellar mass. Stars with an initial mass below about 0.1 $M_\odot$ never reach the temperature required for hydrogen burning. Due to the energy loss by radiation from their surface, these stars slowly contract until the free electron gas becomes degenerate. Since now the pressure does no longer depend on the temperature, the density stays constant and the star stops contracting. During the following evolution of such stars of very low mass the stellar radius remains constant and the star slowly cools, until its surface temperature reaches an equilibrium with the surrounding interstellar radiation field. In the astronomical literature such stars are referred to as "brown dwarfs."

Stars approximately in the range $0.1 < M/M_\odot < 0.4$ can ignite hydrogen burning, but their interior temperature always remains below the temperature of helium ignition. Such stars become degenerate and cool down after the end of the hydrogen-burning phase.[2] In more massive stars the helium core produced by the hydrogen burning eventually reaches the temperature for the thermonuclear conversion of helium into carbon and oxygen. In our Sun this will happen while at the interface between the helium core and the hydrogen-rich outer layers hydrogen is still being converted into helium. Only when the outward moving H-burning shell gets too close to the stellar surface to maintain the H $\rightarrow$ He reaction, the hydrogen-burning shell will be extinguished. By that time a helium burning shell is slowly moving outward. Stars in this post-main sequence shell-burning phases have much larger radii and lower surface temperatures (making them appear redder) than main-sequence stars of the same mass. Therefore, such stars are called "red giants." Because of the extended outer layers (making up most of the volume), the red giants have a much lower mean density than the main-sequence stars. On the other hand, the density of the core, now consisting of relatively heavy elements, is much higher, and (with values of the order 10^9 kg m^{-3}) high enough for the degenerate electrons to dominate the pressure. When finally the helium-shell burning is extinguished, the envelope collapses and the pressure of the degenerate electron gas becomes dominant practically anywhere in the star. Only in a thin outer shell, amounting to less than 1% of the stellar radius, the matter retains the equation of state of the ideal gas. Stars with these properties are called "white dwarfs." The radii of white dwarfs are typically of the order of the Earth's radius, but their masses are similar to that of our Sun. Since the central pressure in white dwarfs increases with the mass and since a higher pressure of degenerate matter requires an adequately higher density, the radii of white dwarfs decrease with increasing mass. (This is qualitatively different from planets and main-sequence stars, where the radii always increase with the mass.)

[2] The main-sequence lifetime of such stars is longer than the age of the universe. Therefore, no post-main sequence stars of this type are observed.

2.1.4 The Evolution and Fate of Massive Stars

As shown by S. Chandrasekhar in 1933, for stars where the pressure is dominated by the degenerate electron gas, hydrostatic equilibrium is possible only below a certain critical mass. This mass limit is due to the fact that at very high densities the velocities of the degenerate electrons approach the velocity of light. In this case a further density increase results mainly in a (special relativistic) mass increase of the electrons, while the pressure can no longer grow fast enough to compensate gravity. The exact value of the resulting upper mass limit of white dwarfs (called Chandrasekhar mass) depends on the chemical composition and on the rotation. For non-rotating white dwarfs it is always below $1.5 M_{\odot}$. Stars which by the end of their final nuclear burning have masses larger than the Chandrasekhar mass cannot become white dwarfs. For this reason (and because of other, more subtle, physical effects, see e.g. Chaps. 31 – 34 of [268]) the evolution of massive stars differs significantly from the evolution of the low-mass stars described in Sect. 2.1.3.

In stars which are too massive to end up as carbon–oxygen white dwarfs, the carbon–oxygen core produced by the helium burning will eventually ignite, producing Ne and Mg. The minimum main-sequence stellar mass for an eventual carbon ignition depends on the exact chemical composition, on the mass loss, and on the rotation, and is not well known. However, it is generally assumed that main-sequence stars of at least up to 4 $M_{\odot}$ (and possibly up to 11 $M_{\odot}$) can loose enough mass during hydrogen and helium burning to eventually end up as white dwarfs with an unprocessed, degenerate carbon–oxygen interior.

For main-sequence stars with initial masses in the range approximately 10–100 $M_{\odot}$, carbon–oxygen burning is followed by additional thermonuclear reactions, which result in atomic nuclei of progressively higher mass. As described for the low-mass stars, several different nuclear reactions can occur simultaneously in shells at different distances from the stellar center. With increasing atomic weight the energy yield of nuclear fusion reactions decreases until no energy is gained any more when the atomic weight of iron (the element with the maximum nuclear binding energy) is reached. When all available nuclear fuel has been spent, a massive star has to use its thermal energy content to keep radiating from its surface. As a result the star can no longer maintain the central temperature and pressure needed to compensate the gravitational forces. Thus, the star looses hydrostatic equilibrium and the star's core starts to collapse.

In stars with masses above about 100 $M_{\odot}$, which have a low initial abundance of heavy elements, hydrostatic equilibrium can be lost already at a much earlier evolutionary stage. In such stars carbon burning temperatures can be reached, where individual photons attain energies which exceed the rest-mass energy of two electrons. In this case photons are converted into electron–positron pairs. Since in massive stars the pressure is dominated by the radiation field, the removal of photons and the cooling by the pair production results in a loss of the hydrostatic equilibrium, followed by a gravitational collapse. This process is known as "pair production instability."

2.1.5 Exploding Stars

In normal stars, energy is produced by means of stable nuclear burning on timescales of millions of years or longer. However, under certain conditions, huge amounts of the energy stored in stars can be liberated on timescales of seconds or minutes. In this case, we observe violent stellar explosions where the stellar brightness can temporarily increase by many orders of magnitude. Stellar explosions have a profound influence on the evolution of galaxies since the ejection matter by such explosions plays an important role for the chemical evolution and the energy balance of the interstellar medium. Moreover, because of their enormous energy output, stellar explosions can be observed over large distances, and consequently over a large range of cosmic epochs. Finally, the occurrence and the relative frequency of the different types of stellar explosions (and their spectra) provide valuable information on the stellar population of a galaxy.

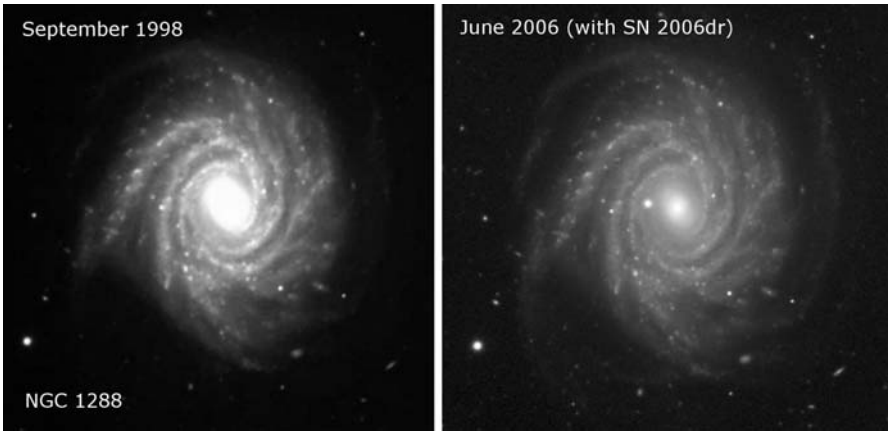

Fig. 2.5 Example of a supernova of Type Ia which exploded in the spiral galaxy NGC 1288. The SN is the bright object to the left of the core of the galaxy in the 2006 image. For a short period in 2006 the SN was about as bright as the rest of the galaxy. (VLT Images, European Southern Observatory)

As stellar explosions are rare—but very luminous—events, we usually observe them at great distances. Often the distances are so large that only the explosion, but not the progenitor star can be detected. Therefore, the ancient astronomers regarded these objects new stars (or "stellae novae" in Latin). For this reason stellar explosions are still called "novae," or — in the case of the more luminous outbursts— "supernovae." Today many different types of novae and supernovae are known. The veritable zoo of different types will not be described here. However, those types of stellar explosions which are particularly important in the context of the evolution of galaxies will be briefly discussed. Most important in this context are the supernovae (SNe), which (for a short period of time) can reach an optical luminosity which is comparable to the total brightness of a small galaxy (Fig. 2.5). From their spec-

tra the SNe are divided into two main types: Supernovae of type I are defined by the absence of hydrogen lines in their spectra. In the spectra of type-II supernovae hydrogen lines are present.

2.1.5.1 Explosions of Single Stars

Type II supernova typically occur at the end of the evolution of massive stars. As pointed out above, after all available nuclear fuel has been exhausted, those massive stars, which during hydrogen and helium burning could not loose a sufficient amount of mass to become white dwarfs, will terminate their evolution in a gravitational collapse. Soon after the collapse of a stellar core has started, the pressure forces normally becomes negligible relative to the gravitational forces, and the stellar density increases on a free-fall timescale. When central densities of the order 10^{15} kg m^{-3} are reached, the atomic nuclei start to decay, neutrons are produced by protons capturing electrons, and a "proto-neutron star" is formed. The increasing density of the degenerate neutron matter results in a rapid pressure built-up. Therefore, at least temporarily, the collapse of the core is stopped and an accretion shock front is formed at the surface of the new, small (radius $\approx$ 30 km) hydrostatic core. During these processes a large amount of neutrinos is generated. Neutrinos are produced in all stellar cores where nuclear reactions take place. However, (because of the extremely small neutrino absorption cross section of ordinary matter) the neutrinos normally escape from the stars without significant losses. Only during the core collapse of a massive star, the matter becomes so dense and hot, that the outer layers of the collapsing stellar core become opaque for the neutrinos produced near the center. As a result, part of the energy of the neutrino flux of the core (of the order 10^{46} W) is deposited in the accreting layers below the shock. If a sufficient fraction of this energy reaches the shock front, the collapse can be stopped and reversed. In this way the initial collapse can result in a highly supersonic ejection of the outer layers of the initial stellar core and of all of the stellar envelope.

The subsequent evolution of the proto-neutron star again depends on the mass. Since neutron stars consist of Fermi-degenerate matter, there also exist a Chandrasekhar mass for neutron stars. However, relative to white dwarfs, neutron stars are much denser. As a result, in neutron stars general-relativistic effects are more important than special-relativistic effects. Solutions of the general-relativistic hydrostatic equation (the Tolman-Oppenheimer-Volkoff equation, e.g., [86]) give upper mass limits of neutron stars, which are lower than the formal Chandrasekhar mass of these objects. But uncertainties concerning the equation of state of neutron matter severely limit the accuracy of these calculations. As a result the exact value of the upper mass limit of neutron stars is not reliably known. Theoretical derivations mostly result in an upper limit near 3 $M_\odot$ (see, e.g., [223, 264]). Observed neutron star masses range between 1 $M_\odot$ and 2.4 $M_\odot$. The very precisely measured masses of the components of double neutron-star binaries all have masses between 1.2 $M_\odot$ and 1.5 $M_\odot$ [95, 223]. Thus the actual upper mass limit of neutron stars is certainly $\geq 1.5 M_\odot$ and probably smaller than 3 $M_\odot$. However, regardless of its exact value, from General Relativity it is clear that an upper mass limit for neutron stars exists.

If, as a result of the mass ejection, the mass of the proto-neutron star formed during the core collapse of a massive star remains below this upper mass limit, a (due to neutrino losses) rapidly cooling neutron star is formed in the center. If the mass ejection cannot prevent the hydrostatic core from growing beyond the critical mass, the matter collapses into a central black hole.

A distant observer of such an event observes a short neutrino flash (lasting a few seconds) which is emitted at the time of the collapse and its partial reversal. The neutrino flash is followed by an optical light outburst, which is produced by the expanding hot bubble of ejecta, and which lasts for a few days or weeks. The light outburst is followed by a slow dimming due to the decay of the unstable radioactive nuclei in the ejecta (typically lasting several months).

The light outburst is produced when the shock wave of the exploding core reaches the original stellar photosphere. Since the surface layers of the evolved massive stars normally still contain hydrogen, explosions resulting from the partial reversal of the collapse of the core of a massive star normally appear as supernovae of type II. However, in very massive stars the core collapse can also occur at an evolutionary phase (such as the Wolf-Rayet phase) when the hydrogen-rich envelope has been completely lost due to stellar winds. In this case the core collapse and explosion of a massive star results in a supernova of type I.[3]

In very massive stars, where the stellar collapse is initiated by the "pair instability" (see Sect. 2.1.4), the sequence of events is somewhat different, since in this case the core (having not proceeded beyond the carbon burning stage) still contains nuclear fuel. As a result, for certain values of the mass and chemical abundance the collapse may lead to a thermonuclear explosion (mainly of oxygen) and a complete disruption of the star. In this case we observe a "pair-instability supernova." For other parameters the pair instability can result in the formation of a massive black hole without significant mass ejection. (For a detailed discussion of this type of supernova see, e.g., [185]; for a detailed discussion of all the various types of explosions of massive stars and their results, see [227, 566].)

Under certain conditions the core-collapse of a massive star can result in the ejection of narrow, highly relativistic jets. If these jets are directed toward us, we observe powerful, short burst of very energetic radiation. This type of supernovae were discovered during studies of the cosmic gamma-ray burst (GRBs), which are flashes of intense gamma radiation with typical durations of the order of seconds. There are two basic types of GRBs: One type is characterized by rather short duration (median ≈ 0.3 s) and a hard X-ray spectrum. The second type has a median duration of about 20 s and a soft spectrum. The observations showed that these "long-soft" or "long-duration" subclass of the GRBs is emitted from core-collapse events of massive stars [566]. These are usually (but not always) accompanied by SNe Type Ic light

[3] More specifically Type Ib, where helium lines are seen in the spectrum (if helium resulting from hydrogen burning is still present at the stellar surface) or Ic, characterized by the absence of hydrogen *and* helium (when at the time of the explosion the mass loss had exposed layers where even the helium had already been burned.)

outbursts. The total amount of energy radiated by GRBs is comparable to that of other supernovae. However, since in the case of GRBs the radiation is beamed into the direction of the relativistic jets, the flux density of the observed electromagnetic radiation from the long-soft GRBs exceeds (typically by about a factor of 10) that of other supernovae. Thus, GRBs are the most intense cosmic radiation sources known at present. In one case (GRB 080319B, see Fig. 7.47 in Sect. 7.4) an optical flux more than 10^3 times higher than that of the brightest observed normal supernova was recorded. In spite of its redshift of $z = 0.937$, this was briefly a naked-eye object (of 5.35 mag).

Although our theoretical understanding of the GRBs is still incomplete, it seems clear that a high angular momentum is a prerequisite of the formation of GRBs (e.g., [349]). It has been suggested that long GRBs occur when massive star cores with an exceptionally high angular momentum start to collapse. A reason for the high angular momentum could be mass transfer in a binary system, or the absence of angular momentum loss by stellar winds during the preceding evolution. GRBs have been found to reside preferentially in galaxies with a low abundance of heavy elements, which could be the reason for weak stellar winds in the massive stars.

Because of their extremely high emitted flux density, and since GRBs (and their host galaxies) can be readily identified by means of gamma-ray surveys, GRB supernovae are potentially important probes for studying the very distant universe.

Part of the matter, which is ejected during a supernova explosion, has experienced nuclear densities and temperatures above the binding energy of atomic nuclei. Therefore, this matter contains many different (stable and unstable) atomic nuclei. In particular, nuclei heavier than iron, which cannot be formed by stable nuclear burning in stellar cores, are readily formed in supernova explosion. Thus, supernova play a decisive role in the chemical enrichment of the universe. Another important effect of the supernovae is their energy input into the interstellar medium of galaxies. Although the most of the energy generated in core-collapse supernova explosions is emitted in the form of neutrinos, a small fraction ($\leq$1%) is emitted in the form of shock waves and relativistic particles, which contribute importantly to the heating of the interstellar gas in galaxies.

2.1.5.2 Explosions in Close Binary Systems

The supernovae described so far, are events occurring during the evolution of massive single stars. Although there are many stars (including our Sun) which have no stellar companions, most stars in our Milky Way galaxy are members of binary or multiple stellar systems. If the distance between the components of a binary system is large compared to the stellar radii (and if this condition remains valid during the evolution of the two components), the binary components evolve like single stars. However, during the post-main-sequence evolution of very close binaries the stellar radius of one (or both) components can increase to a value which exceeds the potential well of the star. In this case, stellar matter from the surface of the "overflowing" star is transferred to the potential well of the companion star. As a result the expanding star will loss mass, while the mass of the companion is increasing.

Because of the strong mass dependence of stellar evolution and since the exchange of matter can modify a star's chemical compositions, mass exchange has a profound influence on the evolution of close binary components. And under certain conditions the mass exchange can lead to stellar explosions.

Of particular importance in the context of binary evolution and cosmology are the explosions classified as supernovae of type Ia. This type of stellar explosions occur in close binary systems where one component, after having lost all its hydrogen-rich surface layers, is transferring helium to the surface of a carbon–oxygen white dwarf. Under certain conditions the accretion of helium onto the white dwarf results in a helium shell on top of the degenerate carbon–oxygen core and in stable helium burning at the bottom of this helium shell. In this way the mass of the degenerate carbon–oxygen core is increasing slowly. When the mass of the core approaches the Chandrasekhar mass for carbon–oxygen white dwarfs, its interior density and temperature becomes so high that carbon burning is ignited. In the interior of a normal star nuclear burning is stable, as any temperature increase results in a pressure increase, leading to an expansion, which in turn reduces the temperature and the reaction rate. In the interior of a degenerate star, where the pressure (essentially) depends on the density only, this negative feedback mechanism is missing. As a result, if nuclear reactions take place in degenerate stellar matter, any accidental increase of the temperature is followed by higher nuclear reaction rates, leading to a further temperature increase, and a thermal runaway. Therefore, nuclear burning in degenerate matter is always explosive and the ignition of the carbon–oxygen core approaching the Chandrasekhar mass leads to a thermonuclear explosion, which completely destroys the star. Since the matter involved in this type of explosion is devoid of hydrogen, its spectrum shows no hydrogen lines. Therefore, the explosions of carbon–oxygen white dwarfs are the most common supernovae of type I and are classified as "type Ia." As in the case of the supernovae of other types, the matter ejected by the type Ia explosions is enriched in heavy nuclei. However, since type Ia supernovae result from (apart from the very high reaction rate) normal nuclear burning, practically all the matter of the original core is transformed into iron and iron-group nuclei. Thus the chemical composition of SN-Ia ejecta differs significantly from that of supernovae type-II ejecta which show a broader distribution of chemical elements. (For details see e.g. [230].)

Practically, all the energy produced in a SN Ia explosion is transformed into kinetic energy and neutrinos, and only a small amount is directly converted into electromagnetic radiation. However, as an intermediate product of the fast nuclear burning a large amount of radioactive ^{56}Ni is produced, which decays into ^{56}Co, and eventually into ^{56}Fe. The heating of the expanding gas by the radioactive decay of ^{56}Ni (and later ^{56}Co) produces most of the observed optical light and causes the observed characteristic light curves of type Ia supernovae.

In contrast to the energy output of the supernova explosions of single stars (which due to the different masses of the progenitor stars and highly different energy production efficiencies covers a large range), the energy output of supernovae of type Ia is relatively uniform, since we have always the same material which explodes (the carbon–oxygen core), and practically always the same amount (the carbon mass

of a carbon–oxygen white dwarf formed by helium burning and reaching its Chandrasekhar mass). Minor variations in the optical energy output of type Ia supernovae (probably related to differences in the ^{56}Ni production) can be determined and corrected using the observed light curve shapes and other directly observed parameters (e.g. [308]). Therefore, like the Cepheids, the supernovae of type Ia can be used as "standard candles" in astrophysics. However, since supernovae are so much brighter than Cepheids, they can be observed even in distant galaxies. As a result, the SNe Ia are the most important "standard light sources" for measuring large distances and for determining the expansion law of the universe (see Sect. 6.2.1).

Another type of stellar explosions occurring in binary systems are the classical novae. Like supernovae of type Ia, novae explosions occur in close binaries where matter is flowing from a normal star onto the surface of a carbon–oxygen white dwarf. However, in the case of the novae the star shedding mass is in an earlier evolutionary stage than in the case of the supernovae type Ia binaries. In the case of classical novae the companion still possesses hydrogen-rich outer layers. Therefore, the matter transferred to the white dwarf is hydrogen-rich, and the mass exchange results in the formation a thin, dense (and at its base degenerate) hydrogen-rich layer on the surface of the white dwarf. As the thickness of this layer grows, the temperature at its bottom increases, until the ignition temperature of hydrogen is reached. Since the burning takes place in degenerate matter, it is (as explained above) unstable and a thermal runaway takes place, until the degeneracy of the hydrogen layer is lifted by the temperature increase. Like in the case of the supernovae Ia, the explosive burning leads to an ejection of matter into the interstellar space. But since only a thin surface layer (containing 10^{-6}–10^{-4} of the mass) of the white dwarf is involved, the energy output and mass ejection of a classical nova is much smaller than in the case of a supernova. On the other hand, nova explosions can occur repeatedly in the same binary system, and the progenitors of novae are more common than progenitors of type Ia supernovae. Therefore, although supernovae clearly dominate the chemical enrichment of the interstellar medium of galaxies, the classical novae also contribute.

2.1.6 Stellar-Mass Black Holes

Supernova explosion can lead to several different types of central remnants. The exploding star either is completely destroyed by the explosion, or a neutron star remains at the center, or a black hole is formed. Neutron stars are the densest known hydrostatic configurations of matter. Thus, if the mass of a stellar core remaining after a supernova explosion exceeds the upper mass limit of neutron stars, the formation of a black hole is unavoidable. The mass of black holes which can be produced in this way is expected to range between the upper mass limit of neutron stars ($\approx$ between $1.5M_\odot$ and $3M_\odot$) and the mass of the most massive stars ($\approx 10^2 M_\odot$). Black holes in this mass range are called "stellar-mass black holes." At present no physical mechanism is known which could produce black holes of lower mass. Therefore, black holes of sub-stellar masses are not expected to exist. Whether black holes

in the mass range $10^2 < M/M_\odot < 10^5$ can possibly form is unclear at present. Observationally (so far) no conclusive evidence for such intermediate-mass black holes could be found. On the other hand, black holes with masses $M > 10^5 M_\odot$ are well known to exist in the cores of galaxies. These "supermassive" black holes are discussed in Sect. 2.2.7.

The characteristic property of a black hole is the existence of a surface from which (except for quantum effects [4]) no matter or light can escape. Inside the critical surface all matter and energy (including light) falls toward the center. For a non-rotating black hole this surface is a sphere with the radius

$$R_s = 2\mathrm{G}M/\mathrm{c}^2 \tag{2.3}$$

where G is the gravitational constant, M the mass of the black hole, and c the velocity of light. R_s is called the "Schwarzschild radius" or "gravitational radius" of the mass M. (For a derivation of Eq. (2.3) and for a comprehensive account of the theory of black holes see, e.g., [86].)

For one solar mass we get R_s = 2.95 km. Thus, the minimum Schwarzschild radius for stellar-mass black holes is about 5 km. Since black holes do not radiate and since their shadowing cross section is very small (of the order R_s^2) isolated stellar-mass black holes are difficult to detect. On the other hand, stellar-mass black holes can be very conspicuous objects, if they are members of a close binary system, where the companion star (either due to an expansion of its outer layers or due to a strong stellar wind) is transferring mass to the black hole. Because of its angular momentum, this matter cannot fall directly into the black hole. Instead, the matter accumulates in a rotating disk. In a Keplerian rotating disk the orbital velocity increases with decreasing distance from the central mass. Therefore, any friction in the Keplerian disk results in an outward transport of angular momentum, an inward transport of matter, and the eventual accretion of the matter by the black hole. The friction also results in a heating of the gas and in thermal radiation from the accretion disks. If magnetic fields are present, the rapid rotation can result in the acceleration of charged particles.

Matter falling or settling into the deep potential well of a black hole gains a large amount of kinetic energy. The binding energy of a particle orbiting a black hole at a radius close to R_s approaches its rest-mass energy. Therefore, if a suitable mechanism can be found to convert the kinetic energy into outward radiation, in principle, all the rest-mass energy of matter falling into a black hole can be emitted [357]. Since (for a non-rotating black hole) circular orbits inside of 3 R_s are unstable, the efficiency of accretion disks is lower than the theoretical limit given above. Matter in the innermost stable orbit around a non-rotating black hole has a fractional binding energy of about 5.7%. A particle which reaches this orbit has to loose 5.7%

[4] In principle, radiation and matter can escape from black holes due to quantum effects (Hawking radiation). However, for stellar-mass black holes the timescale of the energy loss due to Hawking radiation is $>10^{65}$ years.

of its rest-mass energy. Thus, per unit mass an energy up to $0.057c^2$ can be gained from an accretion disk around a non-rotating black hole. For a rotating black holes, even higher efficiencies of up to 42% can be reached (e.g., [53]). But even the $\approx$ 6% efficiency of the accretion disk around a non-rotating black-hole is much higher than that of any other major cosmic energy generation processes. Nuclear burning of hydrogen, e.g., yields only about 0.7% of the rest mass involved. Thus, accretion disks around black holes are extremely efficient power plants.

Because of the high kinetic energy of the matter in the inner black-hole accretion disks, the gas temperatures resulting from the friction processes typically exceed 10^6 K. As a result, such accretion disks radiate mainly at X-ray wavelengths, and stellar-mass black holes are discovered most easily by searching for cosmic X-ray sources. Strong X-ray radiation is also produced by mass-accreting neutron stars, which have radii not much larger than their Schwarzschild radius. But neutron stars cannot be more massive than 3 $M_\odot$. Therefore, in order to identify an X-ray source as a black hole, its mass has to be determined. This is usually possible if the black hole is part of a binary system and if the binary orbital parameters can be derived. For observational data on stellar mass black holes and their observed basic properties see, e.g., [93].

2.2 Galaxies

Apart from bright stars, the most conspicuous feature of the night sky is the bright band of the Milky Way (Fig. 2.6). The bright band traces the symmetry plane of the disk-shaped galaxy in which our solar system resides. Since our Sun is located almost in this symmetry plane, the nearby stars appear isotropically distributed

Fig. 2.6 The Milky Way observed at the Paranal site of the European Southern Observatory in Northern Chile. The narrow light beam pointing toward the Milky Way is a laser beam emitted from one of the VLT unit telescopes and focused to the high atmosphere to generate an artificial star for an adaptive optics system (see Sect. 4.2.1). (Photo by Y. Beletsky, European Southern Observatory)

around us, while the distant, unresolved stars form a circular band around us. An observer located far outside the Milky Way system would see our Galaxy as a thin disk with an approximately ellipsoidal central "bulge." If the distant observer were (like us) situated in the galactic plane, our Galaxy would look approximately like the edge-on spiral galaxy shown in Fig. 2.10d.

In the Milky Way and in nearby extragalactic systems bright stars can be resolved and observed individually. But most other galaxies are far too distant to be resolved into stars. Therefore, distant extragalactic systems appear to us as diffuse "nebulae." According to our present understanding of star formation, stars can form only in (or together with) galaxies. Thus, although stars can be ejected or removed from galaxies by gravitational interactions between stars or with other galaxies (Sect. 2.2.8), all stars begin their lives inside galaxies.

Only three external galaxies (the Andromeda Nebula in the northern sky and the two Magellanic Clouds in the south) are visible to naked-eye observers. But the number of observable galaxies increases dramatically, when telescopes and sensitive light detectors are used. Even with a small telescope many extragalactic objects become visible, and (as illustrated by Figs. 2.7 and 2.8) images of the sky outside the Milky Way band taken with large telescopes and modern detectors show many more galaxies than stars. Like the stars, the individual galaxies are identified either by a catalog number or by their coordinates on the sky. The most widely used catalog for the identification of bright galaxies is the "New General Catalog" (NGC), which was originally published in 1888. As in the case of stars, catalog numbers can be converted to coordinates (and vice versa) using the Internet facilities of astronomical data bases, such as the CDS.

Figure 2.7, and even more clearly Fig. 2.8 show that individual galaxies have conspicuously different morphologies. On the other hand, most galaxy morpholo-

Fig. 2.7 Section of a deep ground-based image of the southern sky. The image covers an area of about 13 square arcmin and contains more than 2500 galaxies. With very few exceptions these galaxies are at distances corresponding to redshifts $z > 0.3$. Thus, the light of these objects was emitted more than 3×10^9 years ago. (ESO VLT image)

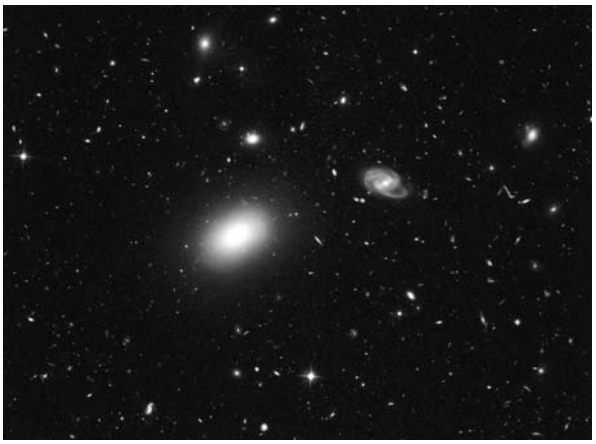

Fig. 2.8 Deep HST image of an about 10 square arcsec area of the northern sky. In contrast to Fig. 2.7, this image also includes nearby galaxies of different types. (Credit: NASA, ESA, and M. Davis, University of California, Berkeley)

gies can be sorted into a few distinct classes, which reflect distinct physical properties and evolutionary histories. Therefore, the following discussion of the galaxies will be started with a definition of the basic morphological classes. This will be followed by a summary of the basic physical properties and the evolution of the different classes. Like the introduction to stellar physics, the summary of the galaxy properties will be restricted to the most basic facts. More detailed accounts of our present knowledge on galaxies can be found, e.g., in [107] and [457].

2.2.1 Morphology, Classification, and Dynamics

Although various different schemes to classify galaxies have been suggested and are being applied, the most widely used classification system is still the one originally suggested by Edwin Hubble [242]. Hubble divided the galaxies into three basic classes, the ellipticals (or E galaxies), the disk galaxies, and the irregular (Irr) galaxies. Since most disk galaxies show spiral patterns of their surface brightness, disk galaxies are also denoted as "spirals" or "spiral galaxies." Many spiral galaxies show inside the spiral patterns elongated stellar concentrations, which are called "bars" (e.g., Fig. 2.10e). Figure 2.9 shows Hubble's original graphic representation of his classification scheme for the regular galaxies. Examples of observed galaxies of different Hubble types are presented in Fig. 2.10. A brief description of the different types is given in the following subsections.

2.2.1.1 Elliptical Galaxies

The elliptical or "spheroidal" galaxies are characterized by approximately elliptical isophote shapes with ellipticities $\epsilon = 1 - b/a$ (b/a being the ratio of the minor and

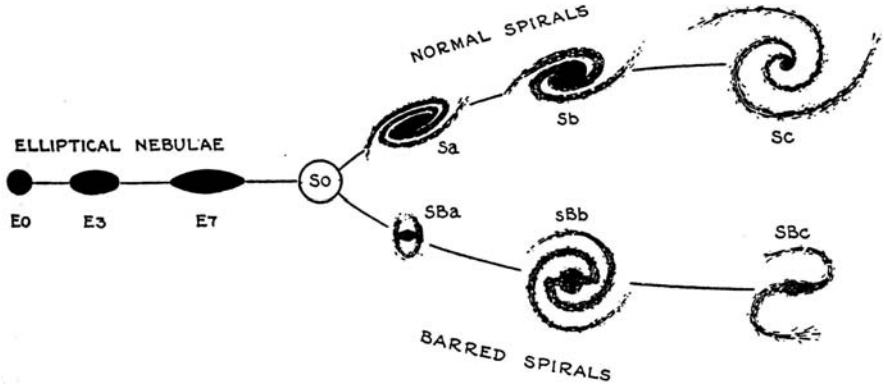

Fig. 2.9 Edwin Hubble's original graphic presentation of his galaxy classification scheme. (Reproduced from [242])

major axis of the isophotes) ranging between 0 and 0.7. E galaxies are divided into subclasses En, where $n = 10\epsilon$. Thus, a perfectly round E galaxy is classified as E0, while a galaxy showing the largest observed ellipticity is classified as E7.[5]

Accurate measurements of the isophote shapes of elliptical galaxies show that E galaxies actually consist of two distinct subtypes [43, 45]. The first of these subtypes show isophotes with (relative to the best-fit ellipses) extensions along the major axes. These galaxies are called "disky ellipticals." The second subtype is characterized by isophotes with box-like deviations from exact ellipses. These galaxies are the "boxy ellipticals." Since the deviations from the elliptical isophotes typically are of the order 1% only, determining the isophote type normally requires an accurate photometry and a Fourier analysis of the isophote shape. Only in extreme cases (as in Fig. 2.11) the isophote type is directly visible on images or on isophote tracings.

Outside the very center, the optical surface brightness I of an elliptical galaxy normally follows (within a few per cent) a so-called "Sérsic law" [467]

$$I(r) = I_e \exp[-\kappa((r/r_e)^{1/n} - 1)] \tag{2.4}$$

where (for a given direction) r is the distance to the center, r_e is the "half-light" or "effective" radius (i.e., the radius inside of which half of the observed total light is emitted), and κ is a constant. The index n ranges between 2 and 8, and depends on the isophote type, being lower than average for disky ellipticals and higher for boxy ellipticals.[6]

Among the basic parameters describing elliptical galaxies are the central velocity dispersion σ of the stars, the apparent mean effective (or "half-light") radius r_e

[5] A different, physically more meaningful, subclassification of the E galaxies has been suggested by Kormendy and Bender [275]. But most current papers still use the original Hubble classes.

[6] In the earlier literature the surface brightness of ellipticals is often fitted using the Sérsic law with a fixed index $n = 4$. This $n = 4$ approximation is known as the "de Vaucouleurs law."

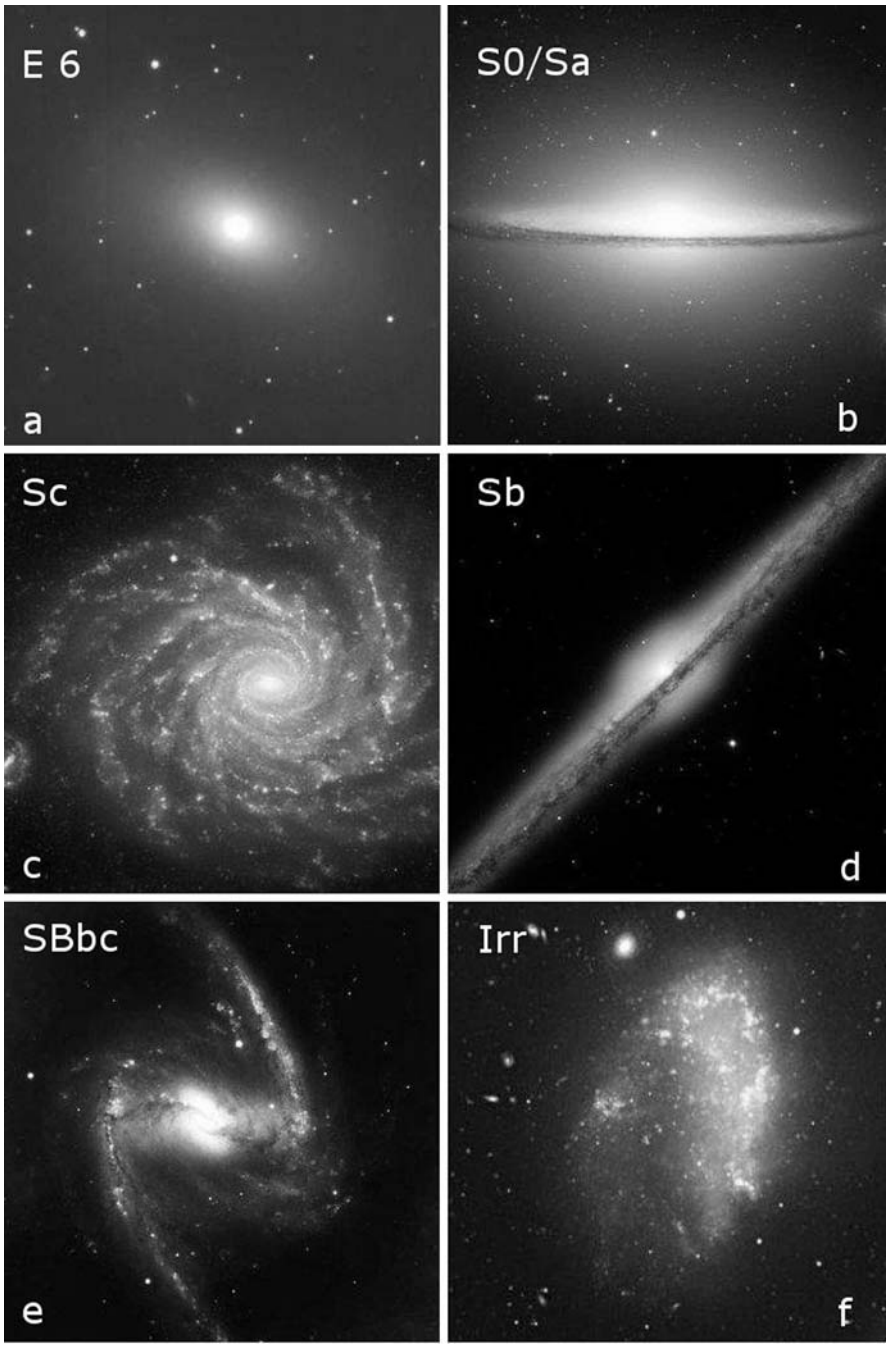

Fig. 2.10 Examples of galaxies of different morphological types and different orientation. The spiral galaxies in the subimages **c** and **e** are viewed along their rotation axis (pole-on), while **b** and **d** present edge-on views of disk galaxies. (Based on images obtained with the ESO VLT)

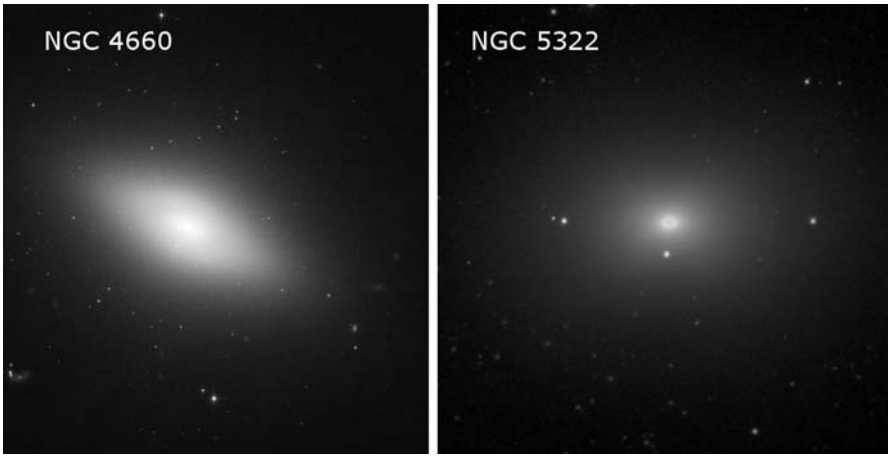

Fig. 2.11 Examples of a "disky" (NGC4660) and a "boxy" (NGC 5322) elliptical galaxy. Quantitative data and isophotes of the two galaxies are given in [44]. (Credit: NASA, ESA, and E. Peng, Beijing University)

(in angular or physical units), and the effective surface brightness $I_e = L/(2\pi r_e^2)$, where L is the Luminosity. The observed values of these three parameters cover a range of values. However, these values are not randomly distributed but follow a scaling relation

$$r_e \propto \sigma^a I_e^b. \tag{2.5}$$

Equation (2.5) obviously defines an inclined 2-D plane in the 3-D coordinate space defined by the logarithms of the three parameters listed above. This plane in the 3-D parameter space is called the "fundamental plane" [138, 145]. The exponents in Eq. (2.5) depend on the wavelength of the observations. Typical values for visual range observations are $a \approx 1.3$ and $b \approx -0.8$. Accurate quantitative data on the fundamental plane relation for local E and S0 galaxies can be found in [256].

All three parameters in Eq. (2.5) are measurable quantities. σ and (in the local universe, where redshifts and the space curvature are negligible) I_e do not depend on the distance, while (for a fixed intrinsic galaxy size) the apparent radius r_e in angular units decreases linearly with distance. Therefore, if the relation between the three parameters is calibrated with nearby galaxies of known distance, it can be used to determine distances of elliptical galaxies. Moreover, the fundamental plane relation is an important tool to probe the physics of spheroidal galaxies (e.g., [431]).

2-D projections of the 3-D fundamental plane scaling relation for elliptical galaxies are known as the "Faber-Jackson relation" $\sigma(L)$ [163] and the "Kormendy relation" $r_e(I_e)$ [274].

2.2.1.2 Disk Galaxies

Disk galaxies always consist of several distinct components. The most conspicuous components are a thin disk and a spheroidal central "bulge."The bulge has properties

which are very similar to those of the E galaxies described above. The disk always is much flatter than the ellipsoidal star distributions of E galaxies. A third major component of the disk galaxies is an extended and nearly spherical halo of stars and star clusters, in which the bulge and the disk are embedded. Sometimes a "thick disk" in the star distribution is observed in addition to the bulge, the normal (thin) disk, and the halo. The "thick disk" has a larger scale height than the thin disk. It consists of stars with a higher vertical velocity dispersion.

The surface brightness distribution of the central bulges of S galaxies follows the Sérsic law observed for ellipticals, while the surface brightness of the disks decreases with good approximation exponentially with distance from the center. (Thus, the disk's surface brightness follows a Sérsic law with $n = 1$.) Moreover, the surface brightness of the disk often shows a spiral pattern. These spiral patterns are due to perturbations in the mass density of the disk, which propagate through the disk as a wave and which have spiral-shaped eigenmodes [317]. The amplitudes of the resulting periodic density variations normally are $<10\%$. However, as explained in Sect. 2.2.6, in spite of this small amplitude, the waves in the density distribution of the stars can efficiently trigger the formation of new stars. Therefore, the spiral density waves result in a rotating spiral pattern of regions of enhanced star formation. Among the newly formed stars are short-lived massive hot and luminous ones. Their light produce the visible spiral arms. The lifetime of these massive stars ($\approx 10^6$ years) is short compared to the rotation period of the density wave pattern (which is of the order 10^8 years). This explains why the arms are so well defined and relatively narrow. Since the hot, blue stars are responsible for the appearance of the spiral arms, the arms are most conspicuous in the ultraviolet and blue light.

As noted already, part of the disk galaxies (including our Milky Way galaxy) show a bar-shaped star concentration in their central region. These galaxies are classified as SB (Fig. 2.9). Following Hubble, the spiral galaxies are subdivided into the classes Sa (SBa), Sb (SBb), and Sc (SBc) (with an extension to Sd by some authors), according to the size and luminosity of the bulge relative to the disk. In Sa galaxies the bulge dominates, in Sc galaxies the bulge is much smaller than the disk. Finally, Hubble introduced the class "S0," which he defined as "intermediate between ellipticals of subclass E7 and Sa or SBa galaxies." S0 galaxies have weak disks, but no spiral arms. Galaxies with types near S0 are sometimes referred to as "lenticulars." But this term is not well defined and it is used with different meanings in the literature.

2.2.1.3 Stellar Orbits and Rotation

In terms of physical parameters, E and S galaxies mainly differ with regard to the amount of rotation, or – more specifically – with regard to the ratio between the rational energy and the total internal kinetic energy of the stellar motions. In all types of galaxies the stars are orbiting the center of gravity of the total mass of the system. In the disks of spirals, essentially all stars revolve in the same sense, in the same plane, and on essentially circular orbits around the center of their host galaxy. Therefore, the disks total kinetic energy is practically identical with the rotational

energy. In the inner few kpc the orbital velocity of the stars and of the interstellar matter of disk galaxies increases about linearly with the distance to the center. After reaching a maximum a few kp from the center, the orbital velocity remains approximately constant over the outer part of the disk. The constant velocity of the outer disk (typically of the order a few 10^2 km s^{-1}) is called the "circular velocity" of the corresponding galaxy.

Rotation also causes the bars in part the disk galaxies. Bars form when the ratio between the rotational energy and the gravitational potential energy (due to all components of the galaxy, including the bulge, the halo, and the dark matter) exceeds a certain critical value [392].

In elliptical galaxies (and in the bulges of S galaxies) the individual stars follow more complex, 3-dimensional, non-circular, and more isotropic orbits. The individual orbits have different planes and may have a different sense of rotation. Thus, although each individual orbit has angular momentum, the sum of the angular momenta of the stars of an E galaxy tends to be small or zero. Therefore, the total internal kinetic energy of an E galaxy is higher than the rotational energy, and there are E galaxies without detectable rotation (and, consequently, without a measurable overall rotational energy). The overall rotation is particularly small or absent in boxy E galaxies, while disky E galaxies always rotate significantly.

Since the disks of spiral galaxies are rotationally flattened, these galaxies are (apart from bars and spiral arms) with good approximation rotationally symmetric. Approximate rotational symmetry is also found in the case of disky E galaxies. In boxy E galaxies the isophote shapes are not related to rotation, but (like in the case of the bars of spiral galaxies) are caused by the non-isotropic distribution of the individual stellar orbits. Therefore, these galaxies normally do not show rotational symmetry, but have triaxial shapes.

As pointed out above, in elliptical galaxies the central velocity dispersion depends on the galaxy size and the effective surface brightness (Eq. 2.5). Thus, σ also depends on the luminosity. Brighter ellipticals tend to have higher velocity dispersions. A similar scaling relation for disk galaxies is the "Tully-Fisher relation" ([529])

$$L(v_{max}) \propto v_{max}^{\alpha} \tag{2.6}$$

where v_{max} is the observed maximal rotational velocity and α is an empirically derived exponent. (For nearby spiral galaxies and luminosities measured in the blue spectral range, the exponent is usually found to be $\alpha \approx 3$.)

2.2.1.4 Other Types

Galaxies not well covered by the Hubble classification are the "dwarf spheroidal" (dSph) galaxies and the "blue compact dwarf" (BCD) galaxies. These two classes represent low-mass galaxies, which look like small ellipticals. But their physical properties (such as surface brightness distribution, rotation, and stellar content) are different. Because of their low luminosity they are observable only in our immediate

extragalactic neighborhood. Since they are not observable at high redshifts, these types will not be discussed here.

Many of the irregular galaxies of the local universe are the result of recent interactions or mergers between galaxies (see Sect. 2.2.8). Other irregulars are small rotationally flattened galaxies which are not massive enough to produce well-defined stellar disks and spiral arms. As there exists transition cases intermediate between these failed S galaxies and true spiral galaxies, special classes (Sdm, Sm, Im,..) have been defined and are sometimes used for such intermediate objects (e.g., [129]).

2.2.2 Size, Luminosity, and Masses

2.2.2.1 Sizes of Galaxies

In the literature, galaxy sizes are listed either as observed angular extents (in arcsec) or as linear sizes (in kpc). If the distance is known, angular sizes can be readily converted into linear sizes. For a comparison of the sizes of galaxies and for a comparison with theoretical models, various different diameters and radii have been defined. Sometimes the total radius r_t of the volume where stars and interstellar matter can be detected is given. A more common measure is the "half-light" or "effective radius" r_e defined in Sect. 2.2.1. Another widely used size parameter is the "Petrosian radius" r_p, which is defined as the radius where the local surface brightness, averaged in an annulus around the center, equals 20% of the mean surface brightness interior to this annulus [411]. In this context also a "Petrosian flux" is defined as the flux observed inside a circle of $2r_p$. In analogy to the normal half-light radius, a "Petrosian half-light radius" is defined as the radius which encloses 50% of the Petrosian flux. For regular galaxies the Petrosian half-light radius is smaller (but of the same order) as the normal half-light radius. r_t always is significantly larger.

An important observational data base on the sizes of galaxies in the local universe has been published by S. Shen and colleagues [474]. Shen et al. based their study on a complete sample of about 140 000 galaxies from the SDSS survey (see 5.2.1). They found that the sizes of all types of galaxies show a log-normal distribution. The two parameters characterizing the distribution are the mean effective radius $\bar{r}_e$ and the scatter around this mean $\sigma_{ln\,r_e}$. Except for very low-mass red galaxies the mean effective radii are found to follow a relation

$$\bar{r}_e \propto M^{\alpha} \tag{2.7}$$

where M is the total stellar mass of the galaxies. For elliptical galaxies the exponent is found to be about $\alpha = 0.55$. In their sample of disk galaxies Shen et al. find $\alpha \approx 0.4$ for galaxies more massive than about $10^{10.6} M_{\odot}$, and $\alpha = 0.15$ for lower total stellar masses. At a given mass the scatter of the effective radii (expressed in kpc) is found to vary between about $\sigma_{ln\,r_e} = 0.3$ for the massive galaxies and $\sigma_{ln\,r_e} = 0.5$ for the low-mass objects, independent of the type. The absolute (Petrosian) mean effective radii listed by Shen et al. range between about 0.1 kpc and about 15 kpc.

2.2.2.2 Luminosities

As in the case of stars, the total radiation power of galaxies (i.e., the energy output per unit time) is called the luminosity. If the radiation output is isotropic, the luminosity can be calculated from the observed radiation flux and the distance. In practice, this task is complicated by the low surface brightness of the outer regions of the galaxies. Therefore, in detail, several different definitions of the galaxy luminosities are in use. Most straightforward are the "total luminosities," i.e., the luminosities derived from integrating the flux over the whole area from which radiation of a given galaxy is received. Since the boundaries of galaxies are not always well defined and since deriving the total luminosities often requires an extrapolation of the observable surface brightness, sometimes "Petrosian luminosities" are used, which can be measured more reliably. Petrosian luminosities correspond to the Petrosian flux defined in the previous subsection. For regular galaxies with a surface brightness distribution according to Eq. (2.4), the Petrosian luminosity equals about 98% of the total luminosity for a S'ersic index $n = 1$ (disk galaxies) and about 80% of the total luminosities for $n = 4$ (elliptical galaxies).

The observed total luminosities of galaxies cover a large range of values. However, the number of observed galaxies Φ as a function of their luminosity can always be approximated by a simple law, which was first suggested by P. Schechter [455], and which is given by the relation

$$\Phi(L) = \Phi^*(L/L^*)^\alpha \exp(-L/L^*) \tag{2.8}$$

where Φ^* is an empirically determined amplitude, α is an empirically derived exponent, and L^* is a characteristic luminosity. For small luminosities ($L \ll L^*$) the Schechter function approaches a power law, while at high luminosities ($L \gg L^*$) the frequency of galaxies drops exponentially. Φ^*, L^*, and the faint-end slope α depend on the observed wavelength range, on the redshift, and on the environment where the galaxies are observed.

In the astronomical literature the luminosity L in the Schechter function is often replaced by the corresponding absolute magnitude M (see Sect. 4.2.4), and Φ is given as the number of galaxies per magnitude range. In this case Eq. (2.8) becomes transformed into:

$$\Phi(M) = (0.4 \ln 10)\Phi^* 10^{0.4(M^*-M)(\alpha+1)} \exp(-10^{0.4(M^*-M)}) \tag{2.9}$$

In the local universe we have for the blue spectral range $\alpha \approx -1$ and $L_B^* \approx 3 \times 10^{10} L_{B\odot}$, where $L_{B\odot}$ is the blue luminosity of the Sun. The total blue luminosity of our milky way galaxy corresponds to approximately L_B^* of the local universe.

More detailed studies show that the Schechter parameters depend on the galaxy type. Figure 2.12 show the (red) luminosity functions and the corresponding Schechter parameters derived separately for the star forming and for the passively evolving galaxies of the local universe. For the star forming galaxies a better fit of the data is obtained by a sum of the two Schechter functions listed in Fig. 2.12.

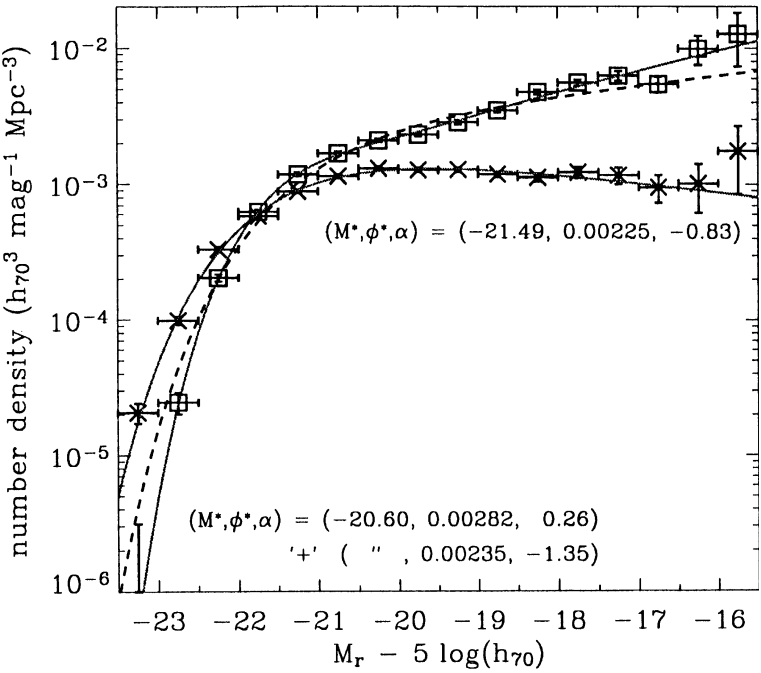

Fig. 2.12 Luminosity functions of the local red-sequence galaxies (*crosses*) and blue galaxies (*squares*) observed in the Sloan Digital Sky Survey. The red-sequence galaxies represent mainly passively evolving systems, while the blue galaxies represent the star-forming galaxies. The *solid lines* correspond to Schechter function fits with the parameters given in the figure. For the blue galaxies a sum of two Schechter functions gives a better fit than a single Schechter function. (The single Schechter function fit is included as a *broken line*). (From Baldry et al. [29])

As shown by Fig. 2.12, relative to the passively evolving systems, the star-forming galaxies are characterized by a steeper faint-end slope α. Thus, at luminosities fainter than about $M_r = -21.5$ there are more star-forming galaxies than passively evolving systems. At the bright end of the distribution the passively evolving elliptical galaxies dominate.

2.2.2.3 Masses

The total mass of a galaxy can in principle be determined from the orbital motions of the stars and of the interstellar gas in the galaxy's gravitational potential. However, for very distant objects velocity measurements have to rely on the Doppler effect, which gives only the line-of-sight (or "radial") components of the velocities. Moreover, often only radial velocities integrated along the line-of-sight through the galaxy can be derived. Deducing mass values from such observations requires model fits. In the case of the circular orbits of the stars in the disks of spiral galaxies, this does not pose problems. However, in elliptical galaxies with their complex stellar orbits such fits are not always unambiguous. Moreover, the observations show that

in addition to the visible matter (observed as stars and interstellar gas) galaxies also contain invisible "dark matter" (cf. Sect. 2.5). The dark matter covers a larger volume than the visible matter which is used to measure the orbital velocities. Therefore, mass derivations based on velocities of the visible matter normally represent only the mass inside the radius where visible matter could be observed. The total mass values listed in the literature cover about the range $10^8 M_\odot$ to $10^{12} M_\odot$ for disk galaxies and $10^7 M_\odot$ to $10^{13} M_\odot$ for elliptical galaxies.

Instead of the total mass, often the "stellar mass" of galaxies is quoted in the literature. In the context of galaxies the "stellar mass" is defined as the total mass of all stars in a galaxy. Since the stellar mass does not include the interstellar matter and the dark matter (Sect. 2.5), it is always smaller than the total (dynamical) mass.

Because of the relation between luminosity and stellar mass, the density of galaxies as a function of their stellar mass can also be fitted with Schechter functions. Like in the case of the luminosities, different fit parameters are found for the mass functions of the star-forming and of passively evolving galaxies [29]. According to Baldry and colleagues [29], the local galaxy mass functions expressed in Schechter parameters give for the passively evolving galaxies a characteristic mass of $1.1 \times 10^{11} M_\odot$ and a faint-end slope $\alpha = -0.87$. For the star-forming galaxies the characteristic mass becomes $4.2 \times 10^{10} M_\odot$ and the faint-end slope is about $\alpha = -1.4$. The number density of passively evolving galaxies is higher for stellar masses, $M_* > 4 \times 10^{10} M_\odot$. At lower masses the star-forming galaxies dominate the local galaxy mass function.

2.2.3 Chemical Composition

According to the current big-bang cosmological models (discussed in Chap. 3), before the formation of the first stars the visible cosmic matter consisted of hydrogen ($\approx 76\%$ of the mass), helium (about 24% of the mass), and traces (of the order 10^{-5}) of deuterium and lithium (e.g., [60]). All other elements were formed later in galaxies. Some galaxies have (due to very low star-formation rates) still compositions close to these initial values. However, in most known local galaxies, the nuclear reactions in stars have reduced the hydrogen fraction and increased the helium fraction. Moreover, a significant amount of heavier nuclei have been produced. The relative abundance (in mass) in the solar atmosphere is approximately 70% hydrogen, 28% helium, and 2% elements heavier than helium. In the astrophysical literature these three mass fractions are often denoted as, respectively, X, Y, and Z, and the chemical composition of the solar atmosphere is referred to as the "solar composition".[7] Thus, the solar composition can be expressed as $X = X_\odot = 0.70$, $Y = Y_\odot = 0.28$, $Z = Z_\odot = 0.02$.

[7] Because of the nuclear reactions in the solar core, the chemical composition of the solar interior and the chemical composition averaged over the whole volume of the Sun differ from that of the solar atmosphere. As a result, the total chemical composition of the Sun differs significantly from the "solar composition," defined as the composition of the solar atmospheric layers.

In the spectra of cool stars a high abundance of heavy elements results in many strong absorption lines of iron and other metals. The strength of these lines depend on the mass fraction of heavy elements (i.e., Z). Therefore, the fraction Z is often referred to as the "metallicity" of an astronomical object. The reliably determined metallicities of local galaxies cover approximately the range $0.02 < Z/Z_{\odot} < 2$.

During the past solar evolution there has been practically no contamination of the solar atmosphere by nuclei produced by reactions in the solar interior. Therefore, the solar atmospheric chemical composition can be regarded as typical for the interstellar matter from which our Sun has been formed. Direct spectroscopic derivations of the heavy-element content of the galactic interstellar gas often give lower abundances than the observations of the solar atmosphere and the solar system objects. This is explained by the fact that in the interstellar space part of the heavy nuclei (but no helium and only a small fraction of the hydrogen) is bound in interstellar dust particles.

With large telescopes and modern spectrographs chemical abundances can be determined not only for galactic stars, but also for individual stars in nearby external galaxies. Moreover, the abundance of certain elements (such as oxygen) can be derived reliably by analyzing and modeling the emission lines of the ionized interstellar gas (see Sect. 6.7). As a result of such studies, major differences of the chemical compositions have been found within the individual morphological classes. However, for all classes the relative abundance of heavy elements was found to correlate with the luminosity of the galaxies. In the local universe massive, luminous galaxies tend to contain a higher fraction of heavy elements than low-luminosity and low-mass galaxies (e.g., [225, 271, 318, 523]). The corresponding relations are known as the "luminosity-metallicity relation" and "mass-metallicity relation." As an example, Fig. 2.13 shows the observed relative abundance of oxygen as a function of the (stellar) mass of galaxies in the local universe.

2.2.4 The Stellar Content of Galaxies

Apart from having different morphologies and rotation, elliptical galaxies and spiral galaxies also differ with regard to their stellar content. Elliptical and S0 galaxies and the bulges of disk galaxies normally contain only old stars with ages >1 Gyr. Since at this age all massive stars with initial masses $>3\ M_{\odot}$ have already exploded or (after mass loss) have become white dwarfs, massive stars are absent in elliptical galaxies. Following W. Baade ([24]), stellar populations which contain only old and low-mass stars are called "Population II," while stellar samples containing young stars (and including massive stars) are denoted as "Population I." Since in Population II systems no new stars are formed and since the only evolutionary aspect is the aging of the existing stars, elliptical and S0 galaxies are often called "passively evolving systems." For historic reasons passively evolving systems often are also called "early-type galaxies," while star-forming systems are referred to as "late-type galaxies." These designations are basically the result of a historic misunderstanding. Therefore, these designations are not used in this book.

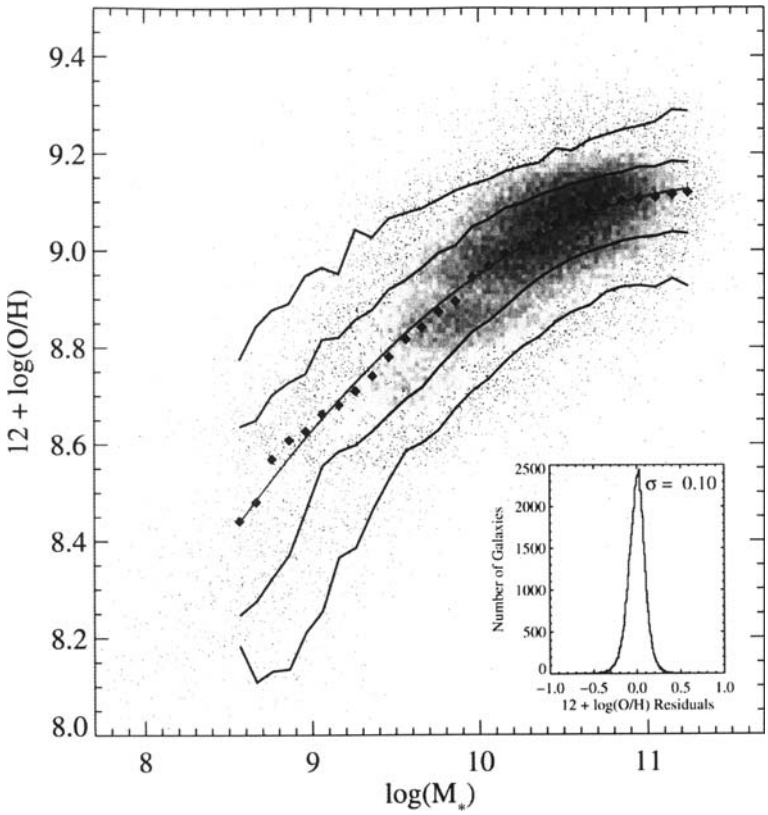

Fig. 2.13 Oxygen abundance relative to hydrogen as a function of the (stellar) mass of galaxies in the local universe. O/H is the number ratio of oxygen and hydrogen atoms. M_* is the total mass of the stars of a galaxy in units of the solar mass. The *small dots* represent individual measurements. The *filled diamonds* represent median values in the corresponding mass bin. The central *solid line* represents a polynomial fit to the data. The outer *solid lines* are the contours that enclose 68% and 95% of the data. The figure is based on the spectra of more than 53 000 galaxies observed during the Sloan Digital Sky Survey (see Sect. 5.2.1). (From Tremonti et al. [523])

During most of their lifetime low-mass stars have relatively low surface temperatures. Therefore, elliptical galaxies and the bulges of spiral galaxies are characterized by red or yellow colors and are bright at near infrared wavelengths. Their spectra are dominated by low-excitation absorption lines produced in the cool atmospheres of low-mass stars (Fig. 2.14). A characteristic feature of these spectra is a break in the continuum energy distributions near the position of the Ca II resonance doublet (at a rest-frame wavelength $\approx$ 400 nm), which is due to a crowding of Fe I and other low-ionization metal lines shortward of this feature.

While elliptical galaxies typically contain Population II stars, a small fraction of the galaxies showing an elliptical morphology has blue colors and contains (also) young and moderately massive stars. These galaxies are called “blue spheroidal galaxies” or BSGs. Detailed studies show that part of the BSGs are precursors of

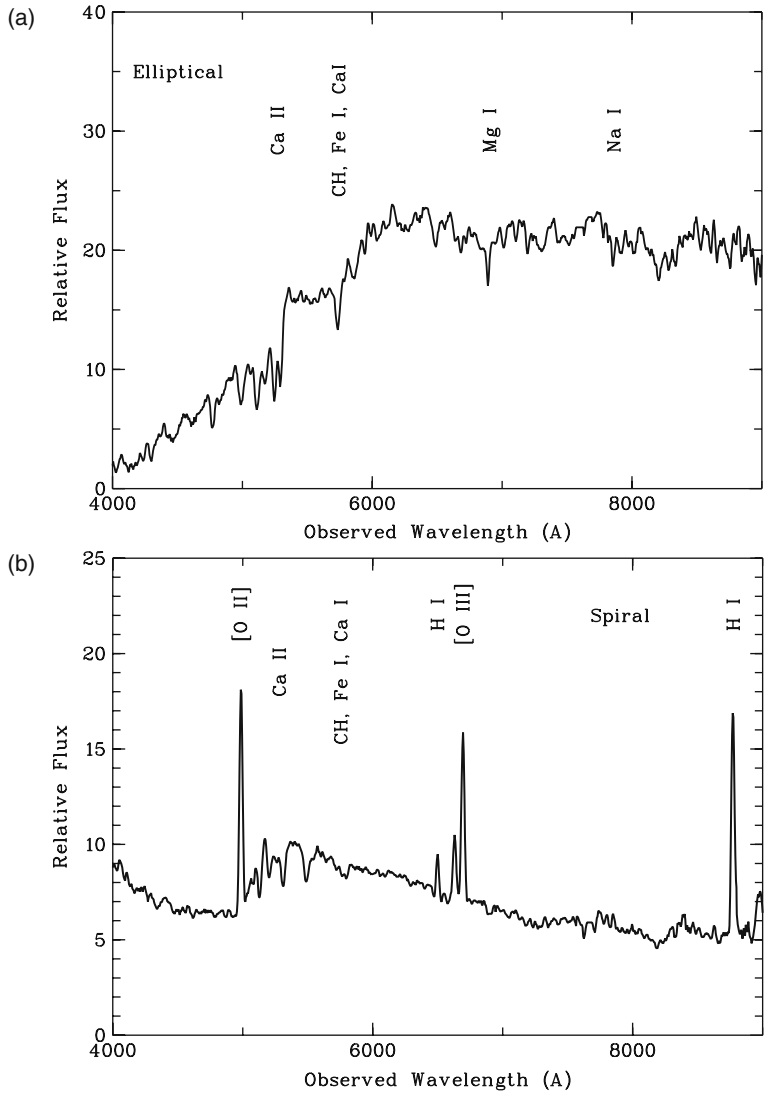

Fig. 2.14 Typical spectra of low-redshift ($z \approx 0.335$) galaxies. The spectrum of the passively evolving elliptical galaxy (*upper panel*) shows mainly low-excitation absorption *lines* produced in the atmospheres of old, low-mass stars. The spectrum of the star-forming spiral galaxy (*lower panel*) is characterized by a bluer continuum and emission *lines* of the ionized interstellar gas

"normal" red elliptical galaxies, where star formation has stopped recently. Other BSGs seem to be low-mass objects related to the disk galaxies and irregular galaxies (for details see, e.g., [222]).

Spiral galaxies contain stars of all ages. In addition to the old stars of the bulges, we find in the disks (also) Population I stars with ages $\ll$1 Gyr. Very young

stellar populations (age <0.1 Gyr) are found in the spiral arms and in the nuclear regions of disk galaxies. Obviously, disk galaxies are characterized by recent or ongoing star formation, and the spiral patterns of the S galaxies trace places of particularly intense recent star formation. Since the stellar luminosities increase strongly with mass (Eq. 2.2), the light from the star formation regions is dominated by the hot, massive young stars. Therefore, the spectra of spiral galaxies have bluer continua and often show interstellar emission lines of the gas ionized by the radiation of the hot young stars (see Fig. 2.14). The blue colors and the emission spectra are particularly conspicuous in Sb and Sc galaxies, while (due to the dominance of the bulges) the spectra of Sa and S0 galaxies may resemble those of ellipticals.

The dominance of the light from hot, young stars also causes the enhanced visibility of the spiral arms in blue and UV images.

The distinct difference between the blue star forming (S, Irr) and red, passively evolving (E, S0) galaxies is also evident from "color-magnitude" diagrams, where the logarithm of a blue/red flux ratio (or "color") of galaxies is plotted as a function of the logarithm of their luminosity expressed in magnitudes (Fig. 2.15). In such diagrams the passively evolving galaxies and the bulges of disk galaxies form a "red sequence" which is distinctly separated (by a "green valley") from the "blue cloud"

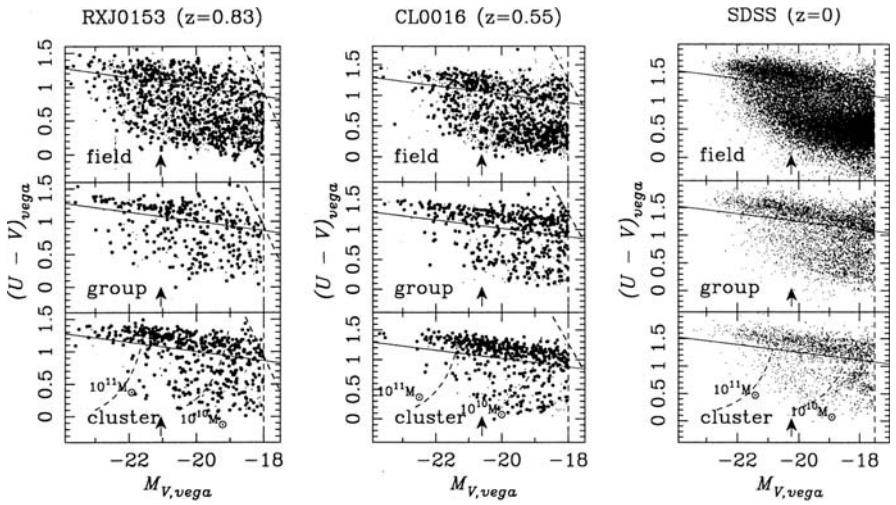

Fig. 2.15 The color index $U - V$ as a function of the luminosity (expressed as absolute visual magnitudes M_V, see Sect. 4.2.4) for galaxies of the Sloan Digital Sky Survey (SDSS, Sect. 5.2.1), representing the local galaxy population, and for moderately redshifted galaxies ($z = 0.55$ and $z = 0.83$, respectively). Redder colors correspond to larger values of $U - V$; higher luminosities correspond to more negative values of M_V. In all three cases observational results for the general field, for small galaxy groups, and for galaxy clusters are given. The "red sequence" formed by the passively evolving galaxies is most conspicuous for the galaxy clusters, but also evident for the galaxy groups and the general field. The "blue cloud" of the star forming galaxies (below the tilted *solid lines*) is less well defined. (From Tanaka et al. [510])

of the star-forming galaxies (e.g., [39, 510]). The red-sequence color-luminosity relation is particularly well defined for the ellipticals of galaxy clusters. This tight relation is caused by an interdependence of the mass, the stellar ages, and the chemical composition of elliptical galaxies (e.g., [193]).

Very young stars are normally found in star clusters or stellar associations (Fig. 2.16). According to the present star formation theory all stars are formed in clusters. However, the internal tidal forces of the galaxies tend to disperse such clusters after some time. Only very massive clusters with a sufficiently deep gravitational well can survive over time intervals comparable to the age of the universe. Such massive clusters reach virial dynamic equilibrium and approximately spherical symmetry. Because of their spherical shape these massive star clusters are called "globular clusters." A prominent example is displayed in Fig. 2.17.

Because of their joint formation, all members of a star clusters have nearly the same age. While the stars of the short-lived small clusters are always young (and thus always belong to the Population I), the population of the long-lived globular clusters depend on the cluster age. Old globular clusters contain only Population II stars. Together with other very old stars, such old globular clusters form the stellar halo of our Milky-Way galaxy.

Since all chemical elements more massive than helium are produced in stars, the first stars must have formed from matter which was practically devoid of heavy elements. Later generation of stars are formed from progressively chemically enriched matter. Therefore, in a given galaxy the initial stellar heavy-element content and the ages of the stars are correlated. However, there are additional parameters which influence the initial chemical composition of stars. In our Milky Way

Fig. 2.16 The young open star cluster NGC 3603. This cluster contains bright hot stars as well as less luminous low-mass stars. It is still surrounded by remnants of its parent interstellar cloud. To reduce the effects of foreground absorption by the interstellar matter this image has been obtained in the near infrared. (VLT image, European Southern Observatory)

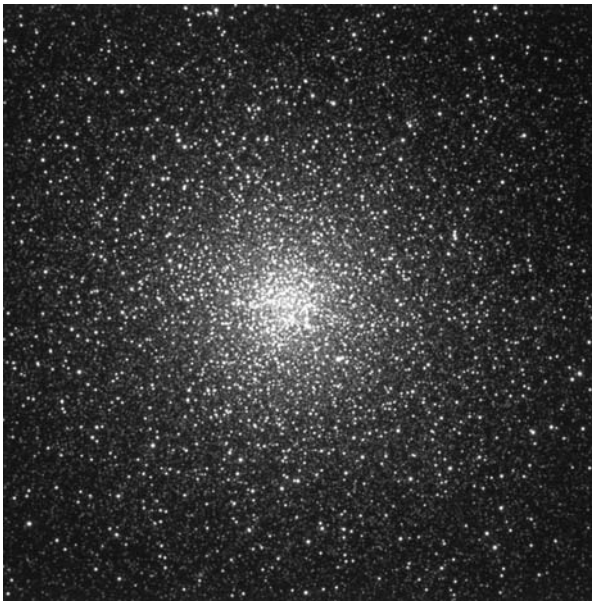

Fig. 2.17 Central part of the globular star cluster 47 Tuc. This star cluster contains only old, low-mass stars. Its morphology is more regular than that of open clusters, and the stars show a more uniform brightness. (VLT image, European Southern Observatory)

galaxy the initial relative abundance of the heavy elements is known to decrease with increasing distance to the galactic center. Finally, as discussed already in Sect. 2.2.3, the average chemical abundances depend on the mass of the galaxies.

2.2.5 The Interstellar Medium of Galaxies

2.2.5.1 Interstellar Gas

Apart from stars, the interstellar gas is the most important constituent of the visible mass of galaxies. All galaxies contain gas in the space between the stars. The main constituents of the gas are always hydrogen and helium. But the properties of the gas vary greatly between the different galaxy types.

In spiral galaxies the interstellar gas has a cloudy structure and is made up of several distinct components. Main components are:

1. clouds of cool molecular gas (consisting mainly of H_2) with temperatures 10–100 K,
2. cold, dense neutral gas, forming extended filaments and sheets, with temperatures 60–100 K,
3. low-density "warm" (several 10^3 K) neutral gas,

4. sheets and clouds of gas, ionized by hot stars, referred to as "H II regions",[8]
5. low-density hot ($T \geq 10^6$ K) ionized gas.

Apart from the dense molecular clouds (were self-gravitation plays a role) the different components are approximately in pressure equilibrium. Thus, the gas densities are about inversely proportional to the temperature. For most components the temperature is maintained by an equilibrium of heating from the local (IR to X-ray) radiation field and the cooling by thermal radiation emitted by atoms, ions, and the associated interstellar dust. The different distinct components correspond to stable states of this equilibrium.

The hot ionized gas, resulting from supernova shock heating, is not in thermal equilibrium, but in the process of losing its internal energy by radiation. However, because of its low density and high ionization, its cooling rate is very low and the cooling timescale is of the order of the universe's age.

The interstellar gas of elliptical galaxies is essentially restricted to the hot ($T \geq 10^6$ K) component. Only few elliptical galaxies also show spectroscopic evidence for small amounts of warm and cold atomic gas. Molecular gas normally is absent in elliptical galaxies.

2.2.5.2 Interstellar Dust

As noted in Sect. 2.1.2, stars lose mass to the interstellar medium by stellar winds. In dense winds of cool stars, in the outflows from novae and supernovae, and, under certain conditions, even in the massive winds of hot stars small (typically μm size) solid particles are formed by condensation. Such "dust particles" can survive (and sometimes even grow) in the interstellar space. Naturally, dust particles are formed mainly from chemical elements and compounds with high condensation and sublimation temperatures. Among the most common constituents of interstellar dust are SiO_2, carbon, carbon compounds, and various metals and metallic minerals. In cool, dense interstellar clouds the dust particles can also grow to "ice mantles" of H_2O, CO_2, or NH_3.

Dust is most conspicuous in the disks of spiral galaxies and in part of the irregular galaxies, where it is associated with cool star-forming clouds. Elliptical galaxies generally contain much less or no dust. In those ellipticals where dust is observed, it is usually concentrated in small central disks. Although the mass fraction of the interstellar dust is generally $\leq$1% of the gas mass, the dust plays an important role for the physics of the cool interstellar medium in disk galaxies. In contrast to atoms and molecules, which can emit and absorb radiation only at wavelengths

[8] Historically the term "H II region" is related to the classification of line spectra. But in astrophysics this designation does not indicate a spectral property. Instead, "H II region" is used to distinguish ionized gas from the neutral interstellar gas. The neutral gas is referred to as "H I gas," following the designation for the line spectra of neutral hydrogen atoms. H II regions emit (due to the recombination of hydrogen) an H I line spectrum. (Ionized hydrogen has no bound states and, consequently, no spectral lines.)

corresponding to transitions of their quantized energy levels, dust grains absorb and emit continuous spectra. Therefore, dust plays an important role for the opacity of the interstellar matter. It is particularly important at temperatures which are too low to result in a thermal excitation of the more abundant atoms and ions. Dust grains which emit a black-body like continuum can emit and cool the gas at any temperature. For this reason dust is most important for cooling the dense, cold interstellar gas. Moreover, because of the dust grains' ability to absorb energy and momentum of atoms and ions, on the surface of dust particles chemical reactions can take place which are rare or impossible in the gas phase.

The actual temperature of the interstellar dust grains is determined by an equilibrium of the energy gain (by the absorption of radiation and by collisions with electrons, atoms, and ions) and the heat loss by radiation. In low-density regions the energy gain is mainly due to the absorption of star light and an equilibrium temperature corresponding to the local effective radiation temperature can be expected. In the solar neighborhood this temperature is about 17 K. The actual local dust temperature is somewhat higher, since the small dust grains cannot radiate well at the wavelength corresponding to this low temperature. Dust near hot stars naturally reaches higher temperatures. However, since dust grains tend to evaporate at temperatures $>10^3$ K, dust always radiates at IR wavelengths. Since only a small fraction of the interstellar dust is found close to hot stars, and since conditions as in our solar neighborhood are more typical, most of the dust radiates at mid and far-infrared wavelengths.

A detailed review of the physical properties and the role of the interstellar dust in astrophysics can be found in [287].

2.2.5.3 Other Components of the Interstellar Medium

Other important components of the interstellar medium are the electromagnetic radiation field (extending from radio wavelengths to the Gamma energy range) and relativistic particles (which for historic reasons are referred to as "cosmic radiation"). Moreover, at least in spiral galaxies, large-scale interstellar magnetic fields are present. The average field strengths are of the order 10^{-10} T and the field lines follow approximately the spiral structure. A detailed review of the properties of the interstellar magnetic fields can be found in [556].

The electromagnetic radiation field and the cosmic-radiation particles are important for heating and ionizing the molecular and atomic interstellar gas. In low-density regions the electromagnetic radiation is the main heating agent, in dense interstellar clouds, which are opaque to visual, UV, and X-ray radiation, cosmic-ray particles dominate the heating of the gas. Together with the magnetic field, the relativistic cosmic ray particles produce synchrotron radiation, which is observed as non-thermal radio emission of spiral galaxies.

Even the neutral interstellar gas normally contains enough charged particles to couple the gas efficiently to the magnetic field. The magnetic field therefore plays an important role for the dynamics of the interstellar medium and for the propagation of waves in interstellar space. In particular, it determines the propagation and structure

of shock waves, since in the low-density ionized medium a momentum transfer via the magnetic field is much more efficient than by means of (rare) particle collisions.

Because of the coupling of the interstellar magnetic field to the gas, the magnetic field strength B varies with the gas density according to $B \propto \rho^{2/3}$, where ρ is the gas density. Since the synchrotron luminosity increases with the magnetic field strength, the strength of the non-thermal radio emission depends on the local gas density as well.

2.2.6 Interstellar Cloud Collapse and Star Formation

2.2.6.1 Triggering Star Formation

Star formation results from the gravitational contraction and collapse of gas clouds. In the present universe practically all stars form from cool and dense gas clouds in disk galaxies and irregular galaxies. Such clouds are normally in hydrostatic equilibrium. The partners of this equilibrium are the internal turbulent and gas pressure, the self-gravitation of the clouds, and the pressure of the surrounding interstellar medium at the cloud surface. This equilibrium can be lost if the density, the cloud mass, or the pressure at the cloud surface is increased. A pressure increase can occur if a pressure wave or shock wave passes through the gas surrounding a cloud. Since the sound velocity inside cool clouds is smaller than in the warm or hot gas, in which the cool clouds are embedded, a pressure wave passes a cool cloud before the cloud can react significantly. Therefore, a passing pressure wave results in an essentially isotropic compression of the clouds.

An important example for star formation triggered by shock fronts is the formation of spiral arms due to the density waves in the disks of spiral galaxies. As pointed out in Sect. 2.2.1, gravitational perturbations cause waves in the number density of stars in galactic disks. These waves form spiral patterns in the stellar density distribution, which rotate with a constant angular velocity. On the other hand, the stars and the interstellar gas are moving on gravitationally bound orbits around the center of the galaxy. Therefore, the gas is moving with a different angular velocity, which depends on the distance to the center. As a result, the matter is streaming relative to the density wave pattern. In most parts of the disks the velocity difference between the gas and the spiral pattern is much higher than the velocity of sound in the gas. When the matter approaches the regions of maximal star density, the gas is accelerated by the gravitational attraction of the density peak. After having passed the density peak, it is decelerated again. As noted in Sect. 2.2.1, the amplitude of spiral density waves is only a few per cent of the mean density. But due to the highly supersonic flow, the resulting modulation of the orbital velocity normally exceeds the velocity of sound in the warm and cool interstellar gas. Thus, the passage of the gas through the spiral pattern results in a shock front, which co-rotates with the density-wave pattern, and which results in a compression of the cool, dense clouds embedded in the gas. Because of the shock compression, the relative density variation in the gas is much higher than that of the amplitude of the variations in

the stellar density distribution. As a result, enhanced star formation occurs in the interstellar matter behind the shock front.

Due to the mechanism outlined above, the spiral patterns can also be traced by the concentration of high-density interstellar clouds and by the non-thermal radio radiation of disk galaxies. Since the radio emission is due to synchrotron radiation, it depends on the density of the relativistic electrons and on the strength of the interstellar magnetic field. Since the field strength as well as the electron density are amplified by the shocks, the radio emission reacts particularly sensitive to the compression of the interstellar plasma.

2.2.6.2 Star Formation in Different Galaxy Types

From the observations it is well established that in normal disk galaxies the star formation rate per unit disk surface Σ_{SFR} correlate with the surface density of the gas Σ_{gas} in the disks. There are two different empirical relations which approximate this correlation about equally well. Most often used is the "Schmidt law" ([456]):

$$\Sigma_{SFR} = (2.5 \pm 0.7) \times 10^{-4} \left(\frac{\Sigma_{gas}}{M_{\odot}\mathrm{pc}^{-2}} \right)^{1.4\pm0.15} M_{\odot}\mathrm{yr}^{-1}\mathrm{kpc}^{-2} \tag{2.10}$$

where Σ denote the disk-averaged surface densities.

The second empirical relation, formulated by R. Kennicutt (see, e.g., [265], where a detailed discussion of the SFR in disks can be found) has the simple form

$$\Sigma_{SFR} = 0.017 \Sigma_{gas} \Omega_g, \tag{2.11}$$

where Ω_g is the circular frequency of the orbital motion of the gas at half of the outer radius of the star-forming disk. Equation (2.11) obviously is equivalent to assuming that during each orbit around the center of the galaxy a fixed fraction (of about $\approx$ 10%) of the available gas mass is converted into stars.

Shocks and the compression of interstellar gas resulting in star formation can also be triggered by the dynamical interaction between galaxies and during the collisions and mergers of gas-rich galaxies. In this case shock fronts may occur in a large fraction of the cool interstellar gas and the resulting total star formation rates can exceed that of normal disk galaxies by orders of magnitude. Galaxies with relative to their mass exceptionally high star-formation rates are called "starburst galaxies." In some of these objects star-formation rates exceeding $10^3 M_{\odot}$ yr^{-1} are observed. In the local universe, starburst galaxies always show evidence of interactions with other galaxies, or of ongoing or recent mergers. Morphologically, starburst galaxies appear irregular or disturbed (e.g., [291]). Their intrinsic spectra are dominated by the blue continua of massive young stars and by strong emission lines of H II regions (Fig. 2.18). A comparison between Figs. 2.14 and 2.18 shows that in the spectra of starburst galaxies the spectral signatures of the hot stars and of the H II regions are even more dominant than in normal spiral galaxies.

In the case of high-redshift galaxies the most accessible spectral region is the redshifted rest-frame-UV range. Therefore, for comparison purposes the UV spectra of starburst galaxies are of particular interest. An example of the UV spectrum of a

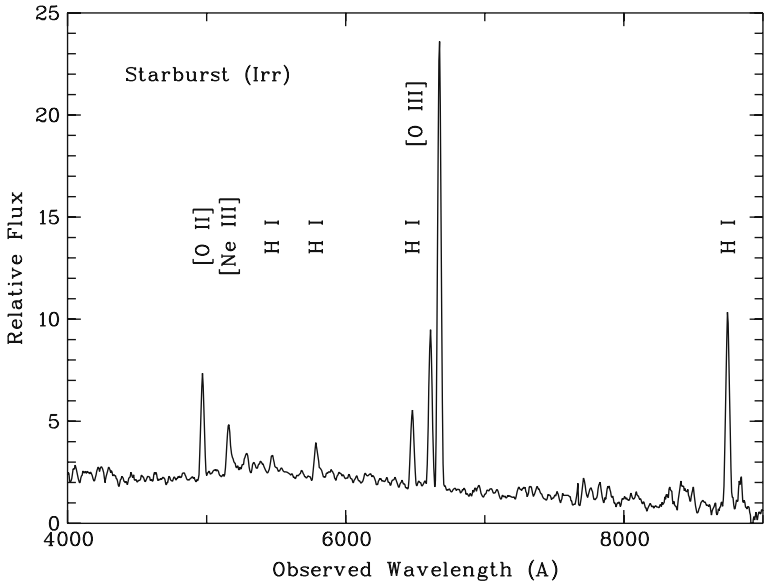

Fig. 2.18 Optical spectrum of the integrated starlight of a low-redshift ($z = 0.333$) starburst galaxy

local starburst galaxy is presented in Fig. 2.19. The figure shows the UV spectrum of the galaxy NGC 4214. This object is a small (Magellanic Cloud like) irregular system with an intense central starburst. Like most high-redshift galaxies, NGC 4214 has a low abundance of heavy chemical elements. As demonstrated by Fig. 2.19, the UV spectra of local starburst galaxies are dominated by the spectral signatures of luminous O-type stars (see Fig. 2.2), which provide much of the observed UV light.

In the local universe high SFRs are always related to the presence of dense, dusty interstellar clouds. Therefore, in many starburst galaxies much of the intrinsic UV flux is internally absorbed by interstellar dust and re-emitted at far-infrared wavelengths. Such dust-obscured starburst galaxies emit no or only little UV flux. Instead we observe relatively uniform dust emission spectra with maxima at wavelengths near 100 μm (Fig. 2.21).

Dust absorption is particularly important in luminous, massive starburst galaxies. Observationally, such objects are known as "Luminous Infrared Galaxies" (LIRGs), or as "Ultraluminous Infrared Galaxies" (ULIRGs) or, if they are detected by the submillimeter tail of the dust emission spectrum, as "submm galaxies." (Fig. 2.21). The massive dust-obscured starburst galaxies were initially discovered by their FIR radiation during the IRAS FIR survey (Sect. 5.2.1). They were later studied in detail with the Infrared Space Observatory (ISO) which was launched in 1995 to observe the infrared sky at wavelengths 2–240 μm. A detailed review of the observed properties of the dust-obscured starburst galaxies and a survey of the literature on these objects can be found in [200]. As a rule, luminous dusty starburst galaxies are found to be related to galaxy mergers or interactions. An example of an IRAS-identified LIRG is given in Fig. 2.20.

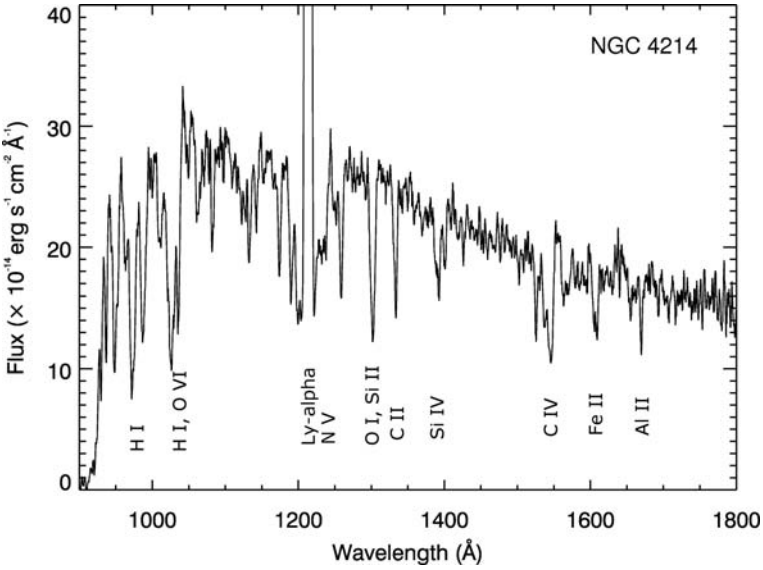

Fig. 2.19 UV spectrum of the integrated starlight of the local starburst galaxy NGC 4214. The broad absorption features (of O VI, N V, S IV, and C IV) are due to stellar winds of luminous hot stars. The *sharper lines* of the lower ionization stages are of interstellar origin. The strong Lyα emission line originates in the outer terrestrial atmosphere (geocorona). This feature obliterates the intrinsic Lyα profile of NGC 4214. (Adapted from Leitherer et al. [311])

Dust-obscured starburst galaxies without detectable optical emission can be identified by their typical FIR-radio energy distributions, which is outlined in Fig. 2.21. Even those local starburst galaxies which can be observed at visual wavelengths often show a significant attenuation of their intrinsic UV flux by dust absorption. Particularly strong dust absorption effects are observed in the Lyα emission lines of starburst galaxies (e.g., [291]).

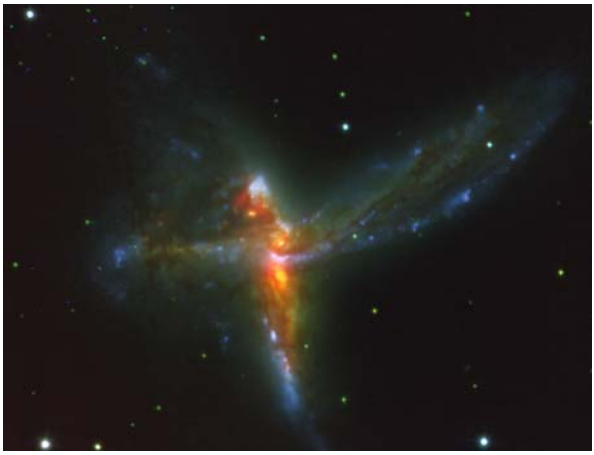

Fig. 2.20 Three-color image of the luminous infrared galaxy IRAS 19115-2124. The image is based on a high-resolution K-band image obtained at the ESO VLT and B and I frames obtained with the HST. A detailed description of this complex merger of galaxies is given in [534]. (Image: European Southern Observatory)

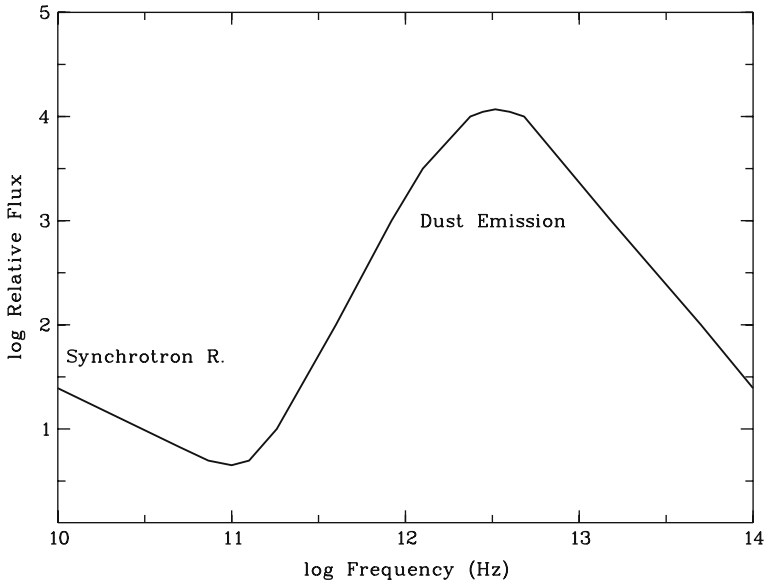

Fig. 2.21 Typical IR and radio spectral energy distribution F_ν of dusty starburst galaxies. The FIR peak of the SED is due to the thermal emission of the dust grains. The radio flux results from ISM synchrotron emission

2.2.6.3 Fragmentation and the Stellar Mass Spectrum

As noted above, cool interstellar clouds normally are in hydrostatic equilibrium. In the presence of a surface pressure, a stable equilibrium is possible for a range of cloud parameters. However, there exists a critical mass, above which no stable equilibrium is possible. In most of the astrophysical literature this critical mass is called the "Jeans mass." The Jeans mass M_J depends on the cloud geometry, the mean cloud density ρ, and on the velocity of sound c_s (where for an ideal gas c_s^2 is proportional to the gas temperature T). For a spherical, isothermal cloud with a constant outside pressure the Jeans mass is given by

$$M_J \approx 1.6 \times c_s^3 G^{-3/2} \rho^{-1/2} \tag{2.12}$$

where G is the gravitational constant.[9] Since clouds less massive than the Jeans mass are stable, gravitational contraction and collapse are most likely to occur in massive interstellar gas clouds.

[9] The numerical factor in Eq. (2.12) depends on the assumed physical and geometrical cloud properties. Since different authors use different (and not always consistent) assumptions and/or different definitions of M_J, different relations and numerical values of M_J are given in the literature. But in all cases the Jeans mass is proportional to c_s^3 and $\rho^{-1/2}$. For a detailed discussion of M_J, see Chap. 26 of [268].

During the collapse of a gas cloud the density is increasing, while the temperature (and thus c_s) does not increase as long as the cloud can cool by radiation. Therefore, as shown by Eq. (2.12), during the collapse the Jeans mass is becoming progressively smaller. In interstellar space the gas temperature normally even decreases with increasing density, since at higher densities there are more atomic collisions, which leads to more excitations of atoms. This, in turn, results in a higher radiative energy loss. Thus, during the collapse of an interstellar gas cloud the temperature and consequently the Jeans mass is decreasing rapidly, and smaller subregions of the collapsing cloud become unstable. As a result, the collapse normally starts with a large cloud with many stellar masses, which during the collapse breaks up into progressively smaller collapsing cloud fragments of eventually stellar mass. This fragmentation process obviously explains why stars tend to form in clusters (Sect. 2.2.4), which (as illustrated by Fig. 2.22) are the most conspicuous tracers of the spiral arms in disk galaxies.

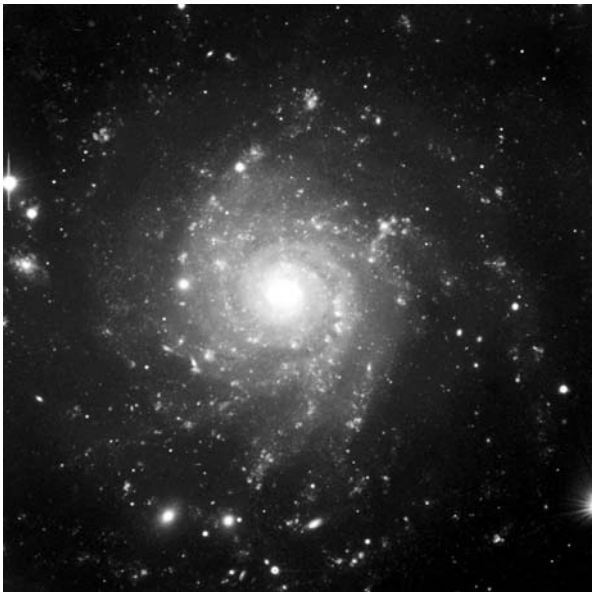

Fig. 2.22 Visual-light image of the spiral galaxy IC 5332. The small spots tracing the spiral arms are open clusters containing luminous young stars, which have been formed during the past $\approx 10^6$ years as a result of a spiral shock wave. (VLT image, obtained at the European Southern Observatory)

During the past decades much progress has been made in numerically modeling star formation. Moreover, significant new insights could be derived from observations of many different stages of the star formation process. On the other hand, because of the non-equilibrium states of the matter of collapsing interstellar clouds, and because of the complex hydrodynamical processes during the fragmentation and final collapse phases, many details of the physics of star formation are still not well understood. Since the formation, evolution, and explosion of stars have major consequences for their host galaxies, the lack of a full understanding of the star formation process is one of the main obstacles to the development of a comprehensive theory of the evolution of galaxies.

2.2.7 Galactic Nuclei

Already the first detailed studies of galaxies indicated special conditions in their central regions. Often bright and variable central point sources are observed, which emit non-thermal radio radiation and X-ray radiation, and which have characteristic emission line spectra, often showing very broad lines. Moreover, in the central regions the surface brightness of elliptical galaxies does not follow the Sérsic law valid for the rest of the galaxy. (For the observed central brightness distributions see, e.g., [301].) Finally, the velocity fields often show peaks near the center. At least in part the special behavior of the galactic nuclei appears to be due to the presence of "supermassive" (typically $10^5 < M/M_\odot < 10^{10}$) black holes. For those galaxies where the mass of the central black holes has been determined from the motions of the stars or of the interstellar gas, the central black-hole masses have been found to be correlated with the spheroidal component of the galaxies (i.e., with the the total stellar mass of the elliptical galaxies, or the bulge mass of the spiral galaxies [276]). On average, about 0.2% of the mass of the spheroidal component is concentrated in the central black hole. For the bulges of 30 nearby galaxies (and for a bulge mass of $5 \times 10^{10} M_\odot$) the median black-hole mass was found to be $(0.14 \pm 0.04)\%$ of the bulge mass [220]. The center of our Milky Way galaxy (bulge mass $\approx 5 \times 10^8 M_\odot$) contains a black hole of about $2.9 \times 10^6 M_\odot$ [460].

As pointed out in Sect. 2.1.6, mass accreting black holes are highly efficient natural power plants, which transform a significant fraction of the rest-mass energy of the infalling matter into radiation. The resulting total luminosity can be expressed as

$$L = \eta \dot{M} c^2 \tag{2.13}$$

where η is an efficiency factor and $\dot{M} = \mathrm{d}M/\mathrm{d}t$ is the mass accretion rate of the black hole. The efficiency factors are predicted to be in a relatively narrow range near $\eta \approx 0.1$ (cf. Sect. 2.1.6). Thus, the order of magnitude of the luminosity is mainly determined by the accretion rate. An upper limit of the accretion rate (or more exactly of the accretion luminosity) can be estimated from the fact that accretion is possible only if the outward acceleration of the matter due to the radiation pressure of the central light source does not exceed the gravitational attraction of the central mass. That the luminosity of self-gravitating objects is limited by radiation pressure effects has first been noted by A. S. Eddington in the context of luminous stars. For (spherical) stars, Newtonian physics, negligible mass loss, the assumption of pure (Thomson) electron scattering opacity, and the stellar luminosity limit are given by the "Eddington luminosity"

$$L_{edd} = 4\pi \mathrm{c} \kappa_{Th}^{-1} G M \approx 4 \times 10^4 (M/M_\odot) L_\odot \tag{2.14}$$

where c is the velocity of light, $\kappa_{Th} = 0.20(1 + X)$ cm^2 g^{-1} is the Thomson mass scattering coefficient, G is the gravitational constant, and M is the stellar mass. The structure and physics of black holes and their light-emitting accretion disks differ significantly from that of stars. Deriving the luminosity limit of supermassive black

holes is more complex and requires a detailed relativistic model of the black-hole-disk system. However, it is generally assumed that the stellar Eddington luminosity (Eq. 2.14) gives at least a correct order-of-magnitude approximation of the limiting luminosity of the black hole sources.

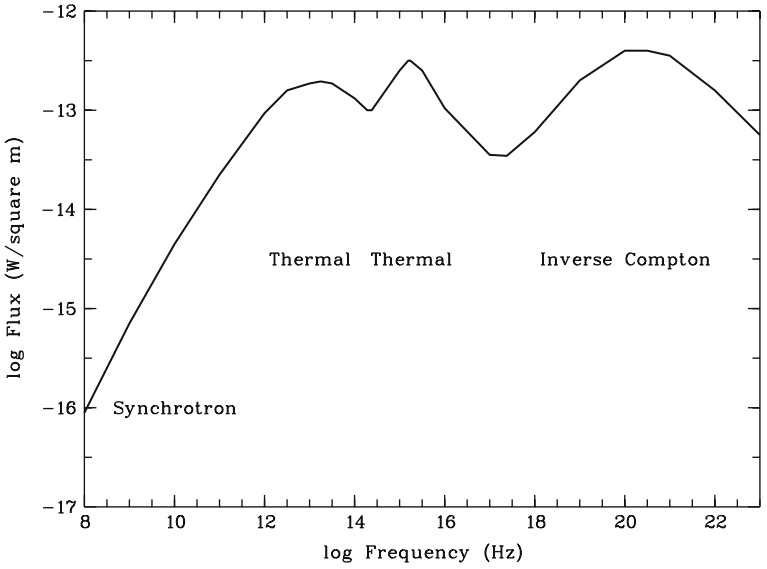

Fig. 2.23 Time-averaged energy distribution $\nu F_\nu(\nu)$ (in W m^{-2}) of the bright QSO 3C273. The distinct spectral peaks are due to thermal dust emission (IR), thermal emission of the inner accretion disk (FUV), and inverse Compton emission (Gamma-rays). (Based on [98] and [113], where a detailed explanation of the spectrum can be found)

There exist stars which have luminosities close to L_{edd} (or, in the case of strong mass loss, even above L_{edd}). But the typical stars forming the spheroidal components of galaxies have luminosities and masses comparable to those of our Sun. Thus, according to Eq. (2.14) the total luminosity of the spheroidal components is by orders of magnitude lower than the Eddington luminosity corresponding to their total mass. On the other hand, if a spheroid harbors a mass-accreting supermassive black hole with a mass near the typical value mentioned above ($\approx 10^{-3}$ of the spheroidal mass), and if this black holes radiates close to L_{Edd} of its mass, according to Eq. (2.14) its total luminosity exceeds the combined energy output of all stars of the host system by a large factor. Such extragalactic systems, where the black hole in the unresolved nucleus outshines the rest of the galaxy, are known as "quasistellar objects" (QSOs) or (since the first known QSOs were strong radio emitters) as "quasistellar radio sources" or "quasars." Less luminous accreting black holes in galactic centers which do not outshine their host galaxies are referred to as "Active Galactic Nuclei" or "AGN." Physically, the QSOs form the bright end of the AGN and thus are part of the AGN population.

Most QSOs and AGN can be identified by their characteristic spectra which consist of three basic components: At radio and Gamma-ray wavelengths the AGN show broad non-thermal continua. Thermal continua are emitted in the UV, X-ray, and IR spectral ranges. Superposed are emission lines produced by gas masses that are ionized by the radiation field of the hot black-hole accretion disk (Figs. 2.23 and 2.24).

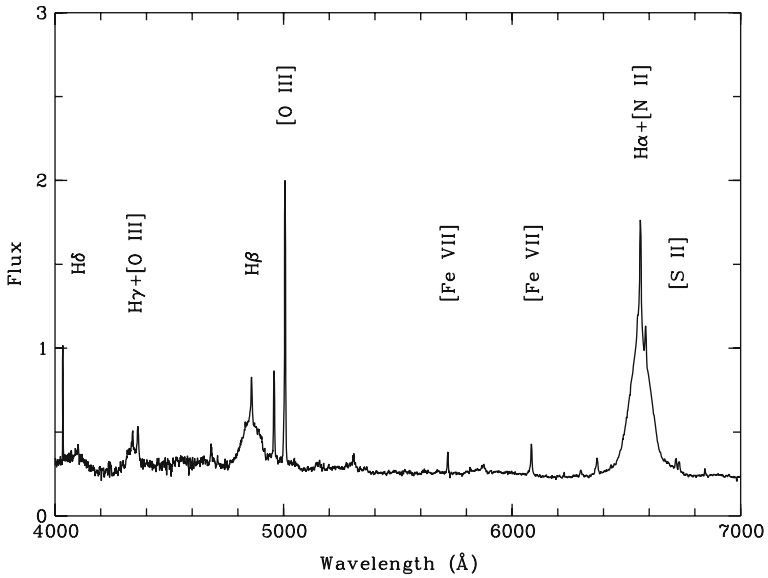

Fig. 2.24 Optical emission *line* spectrum of the AGN Mrk 704. The flux is plotted in units of 10^{-14} erg $cm^{-2}s^{-1}$ $Å^{-1}$. Note the marked differences between the profiles of the individual *lines*. While the permitted H I Balmer lines in each case consist of a broad and a narrow profile component, the forbidden lines ([S II], [N II], [O III], [Fe VII]) show narrow components only

Although central supermassive black holes seem to be a common feature of massive galaxies, only a small fraction of the galaxies appear as QSOs. This indicates that black-hole mass accretion rates producing luminosities close to the Eddington limit are rare events. The small fraction of black holes accreting close to their Eddington limit is not surprising, since galaxies cannot sustain such high mass accretion rates for a long time. At accretion rates close to the Eddington limit all available mass near the galactic centers is exhausted on a short timescale.

An important clue to the origin of the nuclear activity of galaxies are high angular resolution observations of AGN and QSOs. According to these observations galaxies with AGN and QSO, host galaxies often show evidence for recent galaxy interactions or mergers (e.g., Figs. 2.25 and 7.46). Therefore, it is generally assumed that in the case of QSOs and other AGN we see galactic cores where due to galaxy interactions or galaxy mergers (see Sect. 2.2.8) temporarily a large amount of matter is fed into a central black hole. When all the mass in orbits that lead into the galaxy center has been exhausted, and when the disturbed galaxy has reached a new equilibrium

configuration, the mass accretion into the hole ceases, and the black hole becomes inactive. The resulting quiet state, with little or no mass accretion, seems to be the normal state of supermassive black holes in the center of galaxies. An example is the black hole in the center of our Milky Way galaxy which, at present, has a luminosity of only about $5 \times 10^{-11} L_{edd}$. At larger distances such a low activity level could not be detected.

Apart from the extremes of the QSOs and "quiet" nuclei there are many intermediate cases of AGN, where a modest amount of radiation of non-stellar origin is observed and where the nuclear luminosities are comparable to or smaller than those of the underlying galaxies. In all AGN we observe the same basic geometric and physical structure. The central object and power source always is a supermassive black hole, which is surrounded by a hot accretion disk. The disk extends outward from the inner radius of stable orbits (i.e., outward from 3 R_s, where R_s is the Schwarzschild radius of the black hole, see Eq. (2.3). Much further out (at about $10^6 R_s$) the core is surrounded by a torus of cool, dusty molecular gas. Along the rotation axis of the disk and of the torus often narrow, relativistic plasma outflows are observed, which are called "jets." Because of the relativistic bulk velocity of the jet material, radiation emitted by this matter is strongly beamed toward the pole directions of the black-hole disk system. The jets can extend to very large distances (up to about $10^{11} R_s$). Where the jets interact with the ambient interstellar or intergalactic matter, regions of strong radio emission ("radio lobes") are observed. The characteristic broad emission lines of AGN (with typical Doppler widths $10^3 - 10^4$ km s^{-1}, see Fig. 2.24) are produced by small, dense gas clouds at about $10^5 R_s$. The mass of this broad-line region (BLR) is relatively small (of the order $10^2 M_\odot$). But due to the high density of the BLR ($\approx 10^{15}$ atoms m^{-3}) the cooling rate and the energy output of the BLR is very high. The absence of BLR components of the forbidden lines is readily explained by the high density of the BLR gas, which results in a collisional de-excitation of the corresponding energy levels. Beyond the BLR (at about $10^7 R_s$) follows a zone of extended low-density gas clouds producing narrow emission lines. This zone is known as the narrow-line region (NLR). The NLR produces permitted and forbidden lines.

In spite of the common basic structure, the observational properties of individual AGN differ greatly. Depending on the observed properties, the AGN are sorted into three basic classes:

1. Seyfert galaxies [10], which (apart from lower luminosities) have properties like QSOs,
2. radio galaxies (strong radio emitters without the bright optical central point sources which are characteristic of QSOs and Seyferts), and
3. BL Lac objects or "blazars" (with rapidly variable radio, optical, X-ray, and Gamma-ray continuum emission, but with no or only weak optical emission lines).

[10] Named for the American astronomer Carl Seyfert, who in 1943 compiled a first list of such objects.

The Seyfert galaxies and QSOs are further subdivided into type-1 and type-2 objects, depending on the presence or absence of BLR emission.

According to the "unified model" of the AGN, the different AGN classes reflect essentially different inclination angles under which the AGN disk and torus are seen. If an AGN is observed exactly pole-on (i.e., with the line-of sight along its rotation axis), the observed radiation is dominated by the beamed and relativistically amplified (synchrotron) continuum radiation of the jet. Because of the relativistic amplification the jet outshines all other components. Therefore, in this case the AGN appears as a BL Lac object.

At somewhat larger angles between the line-of-sight and the rotation axis the AGN appears as a Seyfert galaxy or QSO of type 1. If the angle between line-of-sight and the rotation axis exceeds a certain value, the accretion disk and the BLR become obscured by the dusty torus. In this case we see a type-2 Seyfert galaxy or a type-2 QSO.

Finally, if we see the system essentially edge-on, all direct optical light from the nucleus is obscured, while the radio radiation from the core and the radio lobes is still visible. In this case we observe a radio galaxy.

In a small fraction of the AGN strong internal absorption leads to spectral signatures which are quite different from those described above (see, e.g., [77, 216]). Some of these QSOs do not show the normally characteristic emission lines but a

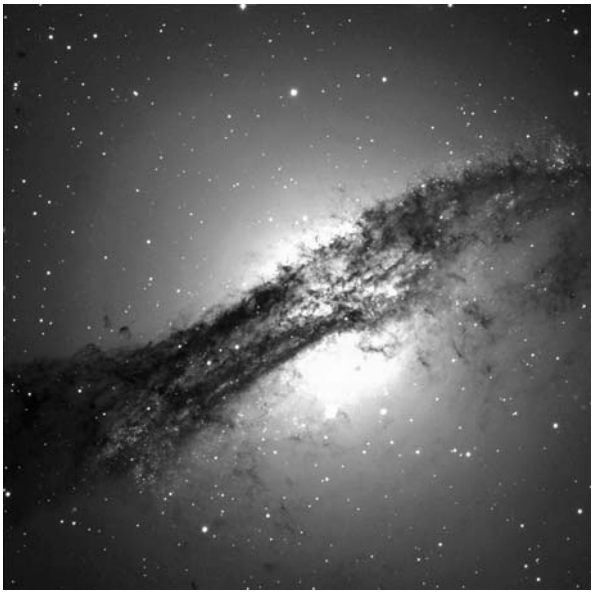

Fig. 2.25 Optical image of the nearby radio galaxy NGC 5128 (Cen A). The highly disturbed morphology of this object is the result of a merger of a massive elliptical galaxy with a dusty, gas-rich galaxy. The dynamical effects of the merger on the gas increase the star formation rate and the mass accretion rate of the central black hole of the elliptical galaxy, which is the likely cause of the observed AGN activity of Cen A. (ESO VLT image)

dense "forest" of overlapping emission and absorption features of iron group elements formed in strong optically thick outflows. In these cases we possibly observe black holes where a period of high mass accretion and energy production has just started. Among these peculiar AGN are high-redshift QSOs with spectra which can easily be mistaken for high-redshift galaxy spectra (see, e.g., [19]).

2.2.8 Interactions Between Galaxies

The distances between stars inside galaxies normally exceed their diameters by factors $> 10^6$. Thus, collisions between stars are extremely rare events. In the case of galaxies the average distances are only about 10–100 times larger than the typical diameters. Therefore, interactions and collisions between galaxies are fairly common.

In the astronomical literature different types of galaxy mergers and interactions have been defined. If a small galaxy is accreted by a much more massive system, the matter of the small galaxy is simply integrated into the massive galaxy, which is not much affected by such a "minor merger." Mergers of galaxies with a comparable mass have more important consequences and may result in fundamental changes of the galaxy properties. Such mergers are called "major mergers." Mergers involving gas-rich galaxies usually are referred to as "wet mergers," while mergers of elliptical galaxies, which do not contain cool interstellar gas, are called "dry mergers." Interactions which do not result in mergers, but which influence the internal state of galaxies are denoted as "harassment."

The statement that stellar collisions are very unlikely also holds when two galaxies collide. Thus, in the case of galaxy collisions the stars pass each other without significant effects for the individual stars. However, while the individual stars remain undisturbed, the overall kinematics and dynamics of the colliding stellar systems are strongly affected. The gravitational field of each star entering and moving through a galaxy accelerates all stars near its path toward its position. As a result, stars move toward the trajectory of the intruding star, and the density of stars temporarily increases in the wake of the star's path. Obviously, this overdensity will always be behind the moving star, and the gravitational field of the temporary overdensity behind the intruding star always attracts and decelerates the moving star. Thus, any star moving through a galaxy loses kinetic energy and momentum by this "dynamic friction" process. During a galaxy–galaxy collision all stars of both galaxies experience dynamic friction. Therefore, the relative motions of the two galaxies are slowed down, and part of the original kinetic energy of the large-scale motion of the galaxies is converted into kinetic energy of random motions of the individual stars. Even if the dynamic friction cannot immediately stop the two galaxies, the galaxies separate again at a reduced relative velocity. This can lead to bound orbits of the two galaxies and to more collisions with further friction effects. In this case the dynamic friction eventually results in the merger of the two colliding galaxies. The chances for a merger are obviously best for head-on collisions of galaxies with small initial relative velocities.

In the case of non-central collisions, and in close encounters not leading to collisions, tidal effects (i.e., differential gravitational fields) are important. Tidal effects can result in strong deformations of galaxies and are the origin of the appearance of many irregular galaxies. Among the most spectacular tidal effects is the stripping of stars from the outer regions of interacting galaxies. These stars often form long tails emanating from the interacting systems (Fig. 2.26). Such tidal tails are rather longlived and can survive for up to several Gyr. The stripped stars may either end up being dispersed in the surrounding intergalactic medium, or eventually fall back into the galaxies.

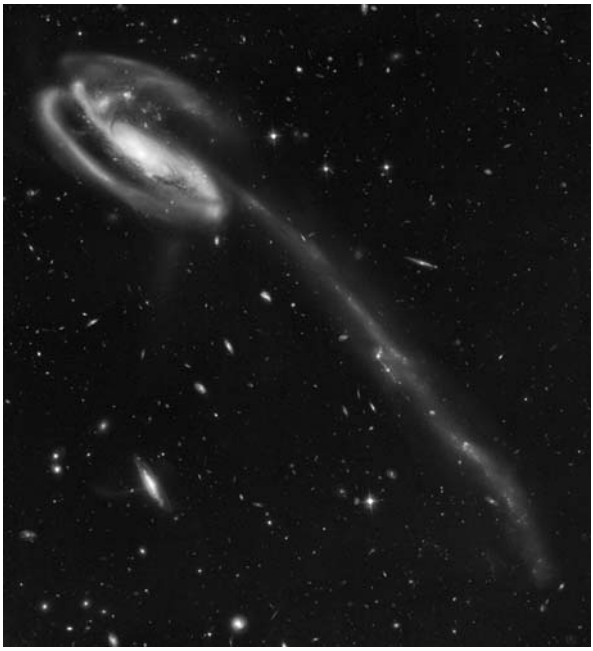

Fig. 2.26 Tidal tails extending from a system of merging galaxies. The tails consist of stars which have been removed from the galaxies by tidal forces during a close encounter preceeding the merger. (HST image. Credit: NASA, H. Ford (JHU), G. Illingworth (UCSC/LO), M. Clampin (STScI), G. Hartig (STScI), the ACS Science Team, and ESA)

As noted above, dynamic friction results in the conversion of kinetic energy of the relative motions of the galaxies into (randomly distributed) orbital kinetic energy of the individual stars. Hence, if spiral galaxies are involved in collisions, their total internal kinetic energy is increased without increasing the rotational energy. Therefore, the disks of such galaxies flare up and the spiral galaxies can acquire the properties of E galaxies. At least part—and possibly all—elliptical galaxies seem to have been formed in this way. In fact, detailed studies of elliptical galaxies often show evidence of a merger origin [45]. On the other hand, interactions of the stellar component of galaxies during merger events cannot explain all observed properties of E galaxies. A possible explanation of these discrepancies are effects of the inter-

stellar gas. If two colliding galaxies contain diffuse interstellar gas or gas clouds (i.e., in the case of "wet mergers"), violent shock fronts are produced in the gas. In these shock fronts the kinetic energy of the gas is dissipated, the gas is heated, and it may be swept out of its host galaxy. Quantitatively, the processes during the mergers of gas rich galaxies are not yet completely understood. Moreover, additional effects (such as gas exhaustion by star formation and the removal of cool gas by the outflows from massive stars, supernovae, and active galactic nuclei) are expected to contribute to the removal of cool gas. The interaction of these different processes cannot yet be modeled well. This may explain why there are still discrepancies between the merger models and some observed properties of the elliptical galaxies.

If merging galaxies contain central black holes, initially a black hole-binary system will be formed in the center of the new galaxy. (Some such systems are observed as binary active nuclei.) However, dynamic friction eventually results in a decay of the binary orbits and a merger of the two black holes.

2.3 Galaxy Clusters and Large-Scale Structure

Like the stars, galaxies are not evenly or randomly distributed in space. Instead, the galaxy distribution has a sponge-like structure. Large (> 10 Mpc) volumes which contain no or only few galaxies (and which are called "voids"), are surrounded by

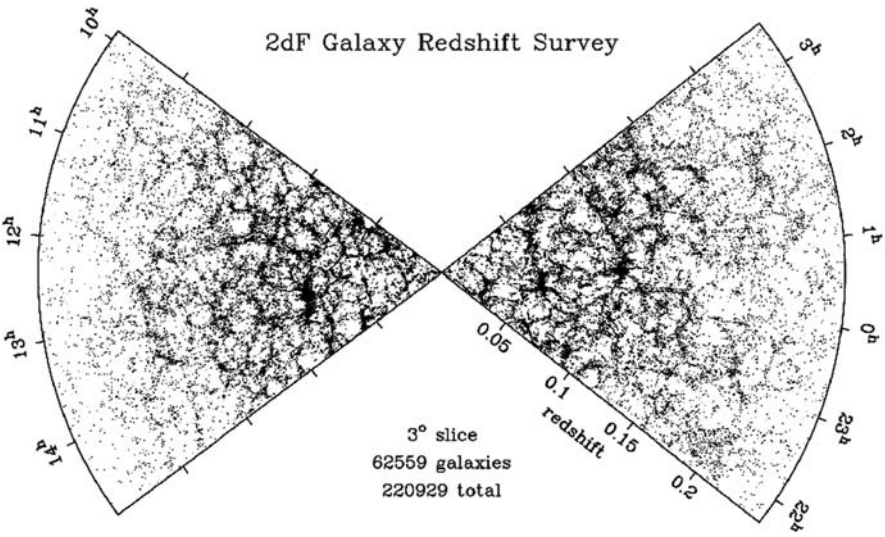

Fig. 2.27 Distribution of the galaxies observed in two slices of the 2dF Galaxy Redshift Survey [106]. The slices have a thickness of 3° in Declination and a total extent of about 10 hours (150°) in R.A. The distance is expressed by the redshift. We are in the center of the image (at $z = 0$). Note the sponge-like large-scale structure, with high-density filaments or sheets surrounding low-density "voids." The density decrease toward the outer edge of the slices is due to the incompleteness of the survey at larger distances. (Credit and copyright: Anglo-Australian Observatory)

sheet-like or filamentary regions of high galaxy densities. This large-scale structure was first seen in galaxy counts based on photographic maps of the northern sky [466]. The filamentary structure is even more obvious in maps of the galaxy distribution based on modern galaxy redshift surveys, such as Fig. 2.27. The regions with the highest galaxy density (usually found at intersections of filaments or sheets of high galaxy density) are gravitationally bound "galaxy groups" or "galaxy clusters" (Fig. 2.28). Galaxy clusters (typically containing between 50 and 10^3 member galaxies) are the largest known gravitationally bound systems in the universe. Their diameters are of the order several Mpc. The (dynamical) masses of clusters and groups cover approximately the range $10^{13} \leq M/M_{\odot} \leq 3 \times 10^{15}$. Most of this mass (>80 %) is dark matter (see Sect. 2.5). About 15% is baryonic matter, mostly in the form of a very hot (T $\geq 10^7$ K) intergalactic gas. The stars of the member galaxies contribute about 3% to the total cluster mass. The intergalactic gas of the clusters has been compressionally heated during the formation of these systems. Because of the low gas density and the practically complete ionization, the (relative inefficient) emission of thermal bremsstrahlung is the main cooling mechanism of the intercluster gas. Therefore, the gas cools very slowly. Nevertheless, since the hot gas forms the bulk of the baryons of the galaxy clusters, these objects are strong emitters of X-rays. As a result, clusters of galaxies are detected most readily by their X-ray emission, and most known galaxy clusters have been discovered during astronomical X-ray surveys.

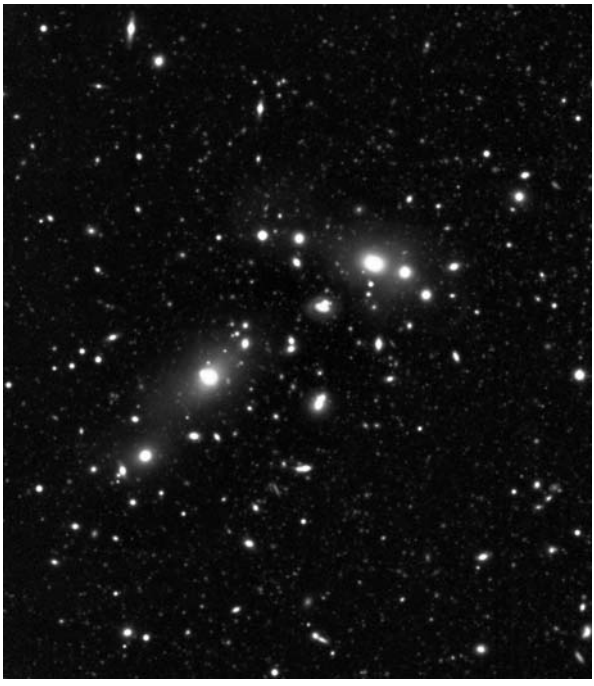

Fig. 2.28 Example of a low-redshift ($z = 0.163$) galaxy cluster. (WFI image, European, Southern Observatory)

The first reliable mass estimate of a galaxy clusters was derived by the Swiss physicist Fritz Zwicky in 1933 [580]. Zwicky determined the total dynamical mass of the Coma galaxy cluster from the observed motions of the galaxies in the cluster's gravitational potential. By comparing the observed total dynamical mass of the Coma cluster with the sum of the masses of its member galaxies, Zwicky discovered (and noted for the first time in [580]) that the visible matter observed as stars and galaxies forms only a small fraction of the observed mass, and that the universe contains a large amount of "dark" matter.

Zwicky's results were later verified with higher precision by many other groups. In addition to the analysis of the galaxy motions, galaxy cluster masses can also be derived from the temperature of the hydrostatically supported hot intergalactic cluster gas. A third method to "weigh" galaxy clusters is the analysis of the gravitational deflection and distortion (often referred to as "gravitational lensing") of light beams from galaxies behind the clusters. As shown and discussed in detail, e.g., in Chap. 13 of [458], and as illustrated in Figs. 2.29 and 2.30 , if the light rays pass sufficiently close to a dense mass concentration, the images of background galaxies become distorted into arc-like structures. From these arc images and the measured redshifts the mass of the deflecting object can be determined.

All three methods mentioned above confirm independently the high dark-matter content of galaxy clusters.

Rich, dense clusters generally consist mainly of elliptical and S0 galaxies. Less dense clusters show a mixture of E and S galaxies. In small groups almost all the stellar mass is concentrated in spiral galaxies. Galaxies (apparently) not belonging to a cluster or a group (the so-called "field galaxies") normally are spirals or small irregular galaxies. Since the high galaxy densities in clusters favors galaxy interactions, the predominance of ellipticals and the absence of spirals in dense clusters obviously supports the suggestion that elliptical galaxies are formed by collisions and mergers.

Fig. 2.29 Example of the amplification and distortion of background objects by the gravitational field of a foreground galaxy (SDSS J120540.43+491029.3). The foreground galaxy has a redshift of $z = 0.215$, the background galaxy distorted into the two inner arcs has a redshift of $z = 0.481$. (HST image. Credit: NASA, ESA, A. Bolton (Harvard-Smithsonian CfA) and the SLACS Team)

Many dense clusters contain a large central E-like galaxy, which is more massive, more extended, and less dense than the normal elliptical galaxies. These galaxies are (for historic reasons) called "cD galaxies" (where the "D" stands for "diffuse"). cDs are probably the result of multiple mergers and the capture of smaller galaxies by a massive central system.

Our Milky Way galaxy is part of a small galaxy group. This "Local Group" consists of three spiral galaxies (the Milky Way galaxy, the Andromeda Nebula, and the galaxy Messier 33) and >30 smaller galaxies (small ellipticals, dSphs, and irregulars).

At least part of the galaxy clusters are members of larger distinct structures, which are called "superclusters." In some cases clearly defined galaxy concentrations with dimensions >100 Mpc have been observed (e.g., [199]). However, in contrast to clusters of galaxies, these very large structures are not gravitationally bound. While galaxy clusters keep their sizes and keep or increase their densities, larger structures partake in the cosmic expansion and are slowly dispersed by the expansion of the universe. On the other hand, such large mass concentration influence the local gravitational potential and the local motions of galaxies relative to the local rest frame. This must be taken into account when analyzing the observed velocities of galaxies relative to our Milky Way galaxy.

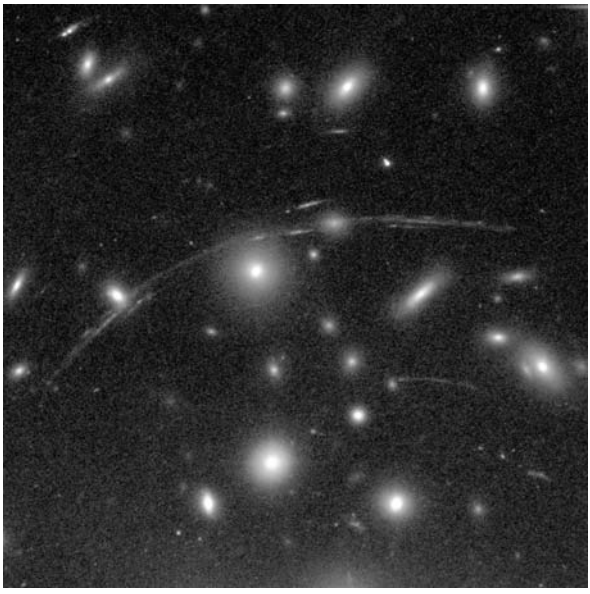

Fig. 2.30 Arc-like distortions of background galaxies by the gravitational field of the galaxy cluster Abell 1689. (HST image. Credit: NASA, N. Benitez (JHU), T. Broadhurst (The Hebrew University), H.Ford (JHU), M. Clampin (STScI), G.Hartig (STScI), G. Illingworth (UCO/Lick Observatory), the ACS Science Team, and ESA)

Although structure in the galaxy distribution is observed even on very large scales, the amplitude of the corresponding density variations averaged over the corresponding scales becomes progressively smaller with increasing scales. Therefore, on very large scales the universe approaches an essentially homogeneous density distribution.

2.4 The Local Intergalactic Medium

The matter density between galaxies is significantly lower than in the interstellar space inside of the galaxies. But, with a baryonic density of the order 0.1 atoms m^{-3} the intergalactic space is not empty. As noted already in Sect. 2.3, clusters of galaxies contain large amounts of hot ($T \geq 10^7$ K) ionized gas. Outside galaxy clusters (i.e., in smaller galaxy groups and in the field) the local intergalactic gas is ionized as well, but its temperature typically is $T \approx 10^5$ K only. From the line profiles of intergalactic absorption lines observed in the near UV it is known that the local intergalactic medium (IGM) is far from uniform and that there exist sheets or filaments of lower temperatures, which are surrounded by hot, low-density gas. Because of its high ionization state, the local IGM gas produces only few and relatively weak spectroscopic signatures. This is illustrated by Fig. 2.31, where the 90–115 nm (EUV) spectrum of a BL Lac object is plotted. Since BL Lac objects have practically featureless continuous spectra, the absorption lines observed in their spectrum must originate in the (galactic) interstellar medium or in the IGM. As shown by Fig. 2.31, most absorption features in the spectra of nearby BL Lacs are of interstellar origin. However, some of the features can be identified as redshifted lines of the Lyman series of H I. For all these H I lines more than one line of the Lyman series has been observed (including the H (1) Lyα lines, which are outside the plotted spectral range). Therefore, the ratios of the wavelengths (which are not changed by the redshift) could be used to identify the lines and the redshift unambiguously.

As shown by the H I lines in BL Lac and QSO spectra, neutral hydrogen atoms are present in the local IGM. But absorption features are seen only at relatively few distinct redshifts, and the H I column density is always low. This confirms that most of the local IGM is completely ionized. Since this ionized component cannot be observed directly, we do not know much about the properties of the local IGM.

Much of the intergalactic gas probably survived since the big bang without much interaction with galaxies, except for radiative effects. But IGM gas located near galaxies (and clusters) can be attracted, captured, and compressed by the gravitational field of these bound configurations. As a result such gas can cool and can eventually be accreted by a galaxy. Matter from galaxies can also be returned to the intergalactic space. Evidence for this feedback is provided by the occurrence of intergalactic absorption lines of elements heavier than helium, such as carbon and silicon. These elements must have been produced in stars inside galaxies, and the corresponding material must have been ejected from galaxies. Mechanisms leading to the ejection of gas from galaxies are stellar winds, supernovae, and galaxy interactions. Although the details are not well understood at present, there is no doubt that an exchange of matter between galaxies and the IGM takes place and influences the IGM as well as the evolution of the galaxies.

All intergalactic matter is immersed in an intergalactic radiation field, which consists of three basic components. Firstly, the intergalactic gas produces a weak diffuse radiation component. More important is the radiation emitted by the stars, the mass-accreting black holes, and the interstellar matter of the galaxies. The most

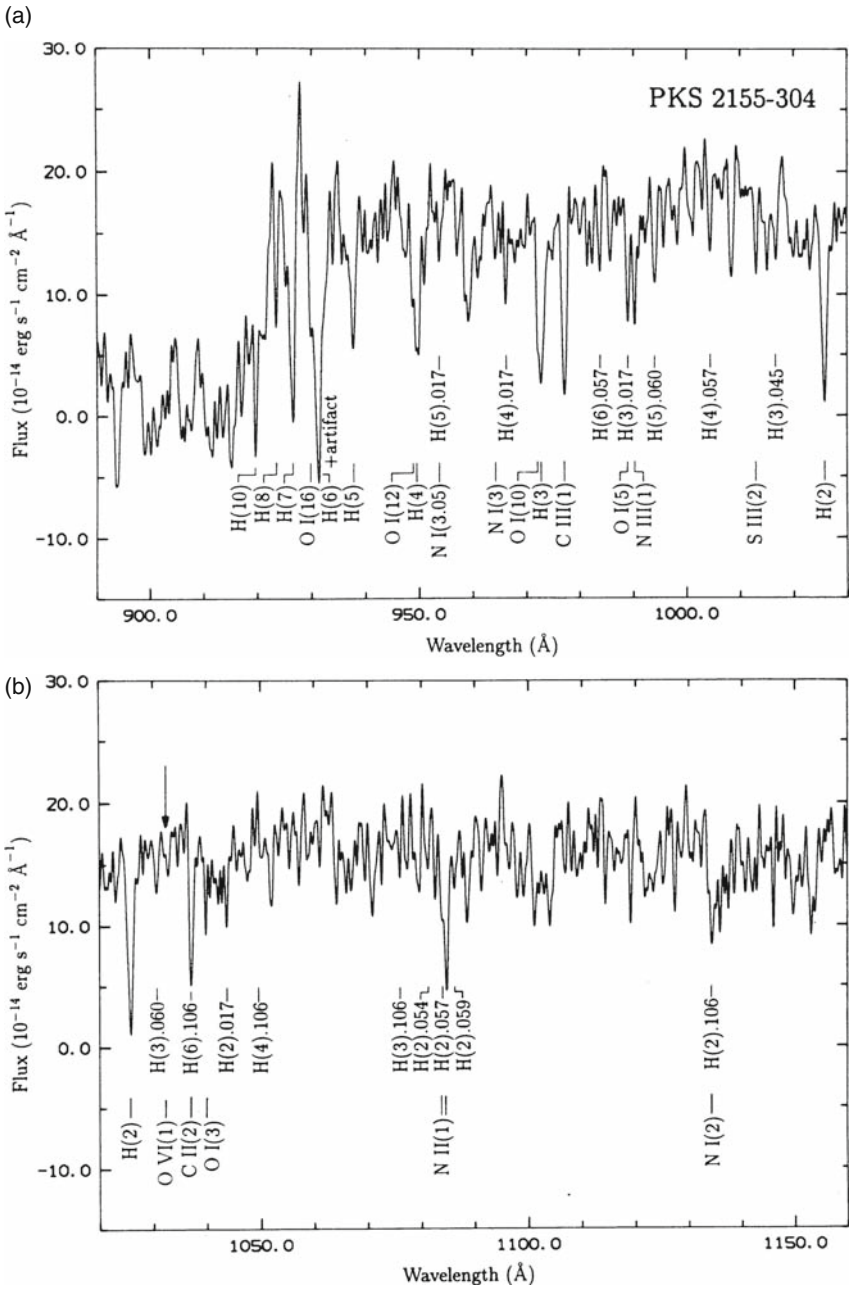

Fig. 2.31 EUV spectrum of the low-redshift BL Lac object PKS 2155-304. All observed absorption lines are produced either in the galactic interstellar medium (lower row of line identifications) or in the intergalactic medium (upper row of identifications). The intergalactic features are due to Lyman-series lines of H I. For these lines the series number (in *parentheses*) and the redshift are given. (From Appenzeller et al. [18])

intense component of the intergalactic radiation field is the Cosmic Microwave Background (CMB), which was emitted by the hot plasma of the early universe (see Sect. 3.2). Today the CMB fills the local intergalactic space uniformly as a black-body radiation of 2.725 ± 0.002 K [340].

2.5 Dark Matter

At present dark matter is defined as matter which shows neither a (with present methods) detectable emission of electromagnetic radiation, nor a detectable absorption of light, nor (except for gravitation) any other detectable interaction with "normal" matter. As already mentioned, the existence of dark matter was first noticed by Fritz Zwicky in 1933 [580] during measurements of the mass of galaxy clusters, where most of the total mass is composed of dark matter (see Sect. 2.3). About 6 years later the Dutch astronomer Jan Hendrik Oort, while analyzing the rotation curve of a spiral galaxy, noted that the outer regions of this galaxy contained much more mass than could be explained by the visible matter [388]. During the following years more such cases were found. However, only about four decades later it was recognized that dark matter is a common constituent of galaxies (see e.g., [162, 483]). Today we know that dark matter is an important part of all galaxies, and that it always dominates the total mass. In fact, galaxies can be defined as accumulations of stars and interstellar matter in the gravitational potential of dark-matter concentrations.

Detailed studies of the gravitational potential of galaxies shows that the dark matter generally fills a volume which is much larger than that delimitated by the stars and by the interstellar matter, and that the dark matter forms extended, massive, approximately spherical halos around the visible galaxies. Obviously, in the present universe the dark-matter distribution is less concentrated than the distribution of the visible or "baryonic" matter.[11]

The dark matter has played a particularly important role during the early history of the universe. Simple numerical estimates show that the mass density of the baryonic matter and the cosmic radiation field has been by far to small to form galaxies (and eventually stars and planets) by gravitational contraction in our expanding universe. Only the dark matter provided enough gravitational attraction to produce bound objects against the dispersive effect of the expansion. Without dark matter neither the present universe, as described in this section, nor we would have been possible.

While the dark matter is invisible and shows (apart from gravitation) no (so far) detectable interaction with other matter or with itself, its distribution can be traced by its gravitational field. From the observed distribution it is clear that dark matter

[11] Since almost all the mass of the normal (visible) matter is due to baryons, in the astronomical literature all matter consisting of atoms and known elementary particles is usually referred to as "baryonic matter."

does not consist of relativistic particles (such as light neutrinos) and that it has been non-relativistic since the time of decoupling from other forms of matter. Dark matter with this property is called "cold dark matter" (CDM). At least 90% of the observed dark matter must be CDM (e.g., [491]).

Dark matter can form gravitationally bound mass concentrations. But, due to the lack of self-interaction, it cannot cool and condense to high densities. This explains why dark matter is most conspicuous on large scales. While dark matter is present everywhere, it has no detectable effect on the scales of star clusters, stars, planetary systems, or black holes. Thus, while galaxies are gravitationally bound mainly by the potential of their dark-matter halos, stars clusters, stars, and black holes are bound by the gravitation of ordinary matter.

In spite of several decades of intense research, the physical nature of the dark matter is still unclear at the time of this writing. It is generally assumed that the dark matter is composed of (thus far) unknown elementary particles. As pointed out above, these particles must be characterized by very weak interactions with other matter and with electromagnetic radiation. Moreover, the particles must have subrelativistic energies. Such properties are to be expected for chargeless, massive particles. Therefore, most physicist regard hypothetic "weakly interacting massive particles" (WIMPS) to be the most likely constituents of the CDM. In the framework of the present standard model of particle physics, no particles with suitable properties are known. But the existence of such particles is suggested by more general theoretical concepts like the supersymmetry (SUSY) models of particle physics ([191]). Apart from WIMPs, candidates for cold dark matter could also be light particles which are generated with non-relativistic energies, such as the hypothetical axions [293]. At present much work is being invested in searching for WIMPS and axions by means of laboratory experiments sensitive to very low-level interactions with normal matter or electromagnetic fields. (For details and a review of the current experiments see [191, 374].) There is hope that these experiments will clarify the elusive physical nature of the dark matter and lead to a physical understanding of this important constituent of our universe within the near future.

2.6 Dark Energy

When formulating his theory of general relativity, Albert Einstein realized that in their original form his field equations excluded a static universe. Since Einstein (as about everybody in 1916) was convinced that the universe was static, he added to his field equations a term which made static solutions possible. Specifically, Einstein introduced the term Λg_{ik} on the right-hand side of the tensor equation Eq. (3.1) (see Sect. 3.1.1), where Λ was assumed to be constant. According to Eq. (3.1) the Λ term determines the space curvature when the mass–energy–momentum tensor T_{ik} vanishes. Therefore, this term can be interpreted as representing the mass–energy–momentum of the empty space.

After astronomical observations proved that the universe was not static (but expanding), Einstein dropped the "cosmological constant" Λ from his equations.

However, precise derivations of the expansion law of the universe, obtained during the past decade, show that the observed expansion of the universe can be explained only if a significant non-zero Λ term is assumed (e.g., [436, 491]). In fact, if interpreted as vacuum energy density, at the present epoch the cosmological constant provides for more than 70% of the total mass-energy density in the universe. This large vacuum energy density, which is required to explain the observed cosmic expansion law, is usually referred to as the "dark energy."

While the energy density of the dark matter described in Sect. 2.5 decreases with the cosmic expansion, with the assumption made by Einstein that Λ is constant, the dark-energy density does not change with time. Therefore, Einstein's field equations can be used to derive the equation of state of the dark energy. Since a constant Λ corresponds to a constant mass-energy density of the dark energy, we get from Eq. (3.7) for the ratio w between the pressure p and the mass-energy density ρc^2 (of the dark energy)

$$w = p/(\rho c^2) = -1 \tag{2.15}$$

A constant density of the dark energy also means that Λ was of little importance in the dense early universe, where the energy densities of the (dark and visible) matter and of the radiation field were much larger than now, and thus much larger than that of the dark matter.

Within the framework of quantum theory the existence of vacuum energy is not unexpected. But presently known physics does not provide reliable predictions on the exact nature and the amount of the energy of empty space. Moreover, according to quantum theory the vacuum energy could have quite different properties than those implicitly assumed by Einstein when introducing his cosmological constant.

A promising theoretical suggestion to explain the dark energy is the assumption of a new type of physical scalar fields. In the present literature such fields are usually referred to as "quintessence" (see, e.g., [406, 555]). In the case of quintessence the energy density of dark matter is not strictly constant but can increase with the cosmic expansion. Quintessence also predicts an equation of state of the dark energy with $w \neq -1$. On the other hand, the observations indicate that w is (at least) very close to -1 (e.g., [491]). Extensive observing programs which presently are being carried out to determine a more exact value of w may decide whether quintessence is a viable candidate for explaining the cosmic dark energy.

Another potential explanation for the observed dark energy follows from theories of physics in more than four dimensions. A dark-energy effect is naturally predicted, e.g., by certain brane-world models (e.g., [127]), although the free parameter space of such models makes an empirical verification difficult.

Since dark energy seems to have become dominant at a relatively recent cosmic epoch, our knowledge on the dark energy is based mainly on observations covering a comparatively small cosmic time interval. Therefore, the concept of dark energy is less well tested than that of the dark matter, which has been confirmed by observations related to many different epochs. Obviously dark energy is a very successful

concept which allows us to explain the observed expansion of the universe with high accuracy. However, in order to ascertain that this concept is the correct one, it will be important to understand the physics of the dark energy.

2.7 A Cosmic Inventory

The preceeding sections gave a brief review of the most important components of the local universe. In Table 2.1 the relative contributions of the different components to the total cosmic mass density are listed. The table is based mainly on a detailed recent discussion of the cosmic mass balance by Fukugita and Peebles [187], who also describe the methods for deriving this data. In addition to the components discussed above, the electromagnetic radiation field and an upper limit of the contribution of the cosmic neutrinos have also been included in Table 2.1. The energy density of the radiation field can be directly measured. The number density of neutrinos can be calculated for the time of their decoupling from other matter (e.g., Chap. 9.4 of [319]). Since that epoch the neutrino field has simply expanded with the rest of the universe. Thus, the present neutrino number density is reliably known. On the other hand, the rest mass of the neutrinos is not well known. Therefore, only upper limits for present mass contribution of the neutrinos can be given.

Table 2.1 Contribution to the present total cosmic mass-energy density of various major components of the universe.

Cosmic component	Mass fraction (in %)	Uncertainty
Dark energy	72	±3
Dark matter	23	±3
Neutrinos	<0.2	
Radiation field	0.005	±0.000004
Intergalactic gas	4.2	±0.3
Stars	0.25	±0.04
Planets	<0.001	
Black holes	0.010	±0.002
Interstellar matter	0.08	±0.01

As shown by Table 2.1, most of the mass in the universe is due to the least well-understood components, like dark energy and dark matter. The baryonic fraction of the mass is dominated by the intergalactic gas. As noted in Sect. 2.4, obtaining accurate information on this component is particularly difficult. The contribution of the most familiar components (stars, planets, interstellar matter) is almost negligible. Obviously, our present knowledge about the universe is very limited and our detailed knowledge is restricted to a very small fraction of its total mass.

Chapter 3
The Past and the Future

As pointed out in Sect. 1.2, observations of nearby astronomical objects show us the universe in its present state, while observations of distant objects allow us to see directly the cosmic past. With increasing redshift we look deeper into the past, and observations at different redshifts allow us to derive the history of our cosmos. Of course, direct astronomical observations of the past are restricted to the time interval during which the cosmos was transparent to light. Fortunately this was the case during more than 99.99% of the age of the universe. Only during the first about 380 000 years (corresponding to redshifts $z > 1090$) the expanding universe was (due to the high free-electron density) opaque to light of all wavelengths.

The redshift z is a directly measurable quantity. However, relating it to the cosmic age and other properties of the epoch at which the observed light was emitted, requires a mathematical model of the cosmic expansion. The following sections summarize the physical basis and the main properties of the expansion models which are currently in use. In addition, this section describes some of the effects of the cosmic expansion on astronomical observations. Finally, our present knowledge and assumptions on the formation of structure in the expanding universe and on the resulting formation of the first stars and galaxies are discussed. Details on these topics can be found in the textbooks on cosmology, such as [403], [405], and [457].

Apart from describing the past, models of the cosmic expansion can also be used to predict the future of the universe. Therefore, the present chapter will end with a brief outlook to our cosmic future.

3.1 The Expanding Universe

3.1.1 Basic Equations and Parameters

3.1.1.1 Basic Equations

On very large scales, gravitation is the governing force in the universe. In the local universe the observed gravitational phenomena are in excellent agreement with Einstein's theory of General Relativity (GR). In spite of many attempts, so far, no

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_3,

conclusive deviations from the predictions of GR could be detected. Therefore, most current cosmological models are based on Einstein's (GR) theory.

The basic equation of Einstein's theory is a tensor equation relating the curvature of space-time (characterized by the metric tensor g_{ik}, the corresponding Ricci tensor R_{ik}, and the contraction R of R_{ik}) to the mass–energy–momentum tensor T_{ik} and the cosmological constant Λ. Einstein's tensor equation can be written as

$$R_{ik} - \frac{1}{2} g_{ik} R = \frac{8\pi \mathrm{G}}{c^4} T_{ik} + \Lambda g_{ik} \tag{3.1}$$

where G is the gravitational constant and c the velocity of light. Because space-time has four dimensions, the tensor Eq. (3.1) corresponds to 16 scalar equations. Due to the symmetry of all tensors involved, only 10 of these equations are independent. These 10 independent scaler equations are referred to as Einstein's field equations. Since the Ricci tensor is a complex, nonlinear expression of the g_{ik} and their first and second derivatives, the field equations form a set of 10 nonlinear partial differential equations.

3.1.1.2 Distances in Space and Time

As described in Sect. 2.3, when averaged over sufficiently large scales, the present universe is highly homogeneous. Moreover, the cosmic microwave background, which was emitted when the universe was about 380 000 years old is (apart from the tiny intensity fluctuations described in Sect. 3.2.3) essentially isotropic. This shows that at early times the universe was very nearly homogeneous, even on small scales. Therefore, for modeling the overall evolution of the universe, the cosmos can with good approximation be assumed to be homogeneous and isotropic. The homogeneity implies that the curvature of space-time is constant in space (being everywhere either positive, negative, or zero) and that a universal cosmic time can be defined. Under these assumptions the differential of distances in space-time can be expressed by the Robertson-Walker line element

$$\mathrm{d}s^2 = c^2 \mathrm{d}t^2 - a^2(t) \left[\frac{\mathrm{d}r^2}{1 - kr^2} + r^2 \left(\mathrm{d}\theta^2 + \sin^2 \theta \mathrm{d}\phi^2 \right) \right] \tag{3.2}$$

where r, θ, and ϕ are polar coordinates in the 3-dimensional space, t is the universal cosmic time, and a is a (properly normalized) time-dependent scale factor. k depends on the curvature, being 0, $+1$, or -1 for, respectively, flat space, positively curved space, and negatively curved space. The scale factor a has the dimension of a length, while the coordinate r, as defined above, is dimensionless. With the above definitions the surface of a sphere with radius r in 3-dimensional space has the surface $4\pi a^2 r^2$, a great circle on the 3-dimensional sphere has the circumference of $2a\pi r$, and the angular extent of an object of fixed projected linear size decreases $\propto r^{-1}$. Therefore, r is called the "angular diameter distance coordinate"or "angular

size distance coordinate." For $k = +1$ the volume of the 3-dimensional space is finite and $r < 1$.

If we assume that we are at the origin of the coordinate system ($r = 0$), r can be used to describe radial (i.e., $d\theta = d\phi = 0$) distances. For distant objects which are locally at rest (i.e., objects which follow the universal expansion without moving relative to their surroundings), the coordinates r, θ, ϕ, and distances expressed in these coordinates do not change with time or redshift. Therefore, these coordinates are called "comoving coordinates" and distances expressed in r are called "comoving angular size distances."

Using the substitution $r_D = \int_0^r (1 - kr^2)^{-1/2} \mathrm{d}r$ Eq. (3.2) can also be written

$$\mathrm{d}s^2 = c^2 \mathrm{d}t^2 - a^2(t) \left[\mathrm{d}r_D^2 + r^2 \left(\mathrm{d}\theta^2 + \sin^2 \theta \mathrm{d}\phi^2 \right) \right] \tag{3.3}$$

(where $r = \sin r_D$ for $k = +1$, $r = \sinh r_D$ for $k = -1$, and $r = r_D$ for $k = 0$)

From 3.3 it is clear that r_D is directly proportional to the radial distance, and that r_D (like r) is independent of time and redshift. Therefore, r_D is called the "comoving radial distance coordinate." Distances based on r_D are "comoving radial distances." Correspondingly, comoving volumes can be defined.

For radial distances (where $\mathrm{d}\phi = \mathrm{d}\theta = 0$) along light rays (characterized by $\mathrm{d}s = 0$) we get from Eq. (3.3): $c^2 \mathrm{d}t^2 = a^2 \mathrm{d}r_D^2$. Thus, for r_D we obtain

$$r_D = \int_t^{t_0} \frac{c \mathrm{d}t}{a(t)} \tag{3.4}$$

where t is the cosmic time at the epoch of the light emission and t_0 is the present-day cosmic time.

3.1.1.3 Friedmann's Equations

By inserting the metric tensor corresponding to (3.2) into Einstein's field equations and by approximating the mass-energy-momentum tensor of the cosmic matter and radiation in (3.1) by that of an ideal fluid with density ρ and pressure p, and defining $\dot{a} = \mathrm{d}a/\mathrm{d}t$, one obtains a new set of differential equations

$$\frac{\dot{a}^2(t)}{a^2(t)} = \frac{8\pi G}{3} \rho(t) + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2(t)} \tag{3.5}$$

and

$$\ddot{a}(t) = \frac{\mathrm{d}^2 a}{\mathrm{d}t^2} = -\frac{4\pi G}{3} a(t) \left(\rho + 3pc^{-2} \right) + \frac{a(t)\Lambda c^2}{3} \tag{3.6}$$

and by combining the two preceding equations:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} a(t) = -3 \left[\rho(t) + p(t) c^{-2} \right] \frac{\mathrm{d}a}{\mathrm{d}t} \tag{3.7}$$

Within the approximations made above, the solutions of these ordinary differential equations fully describe the structure and evolution of our cosmos. Equation (3.5) (and Eq. (3.6) in a less general form) were first formulated by the Russian mathematician A. A. Friedmann ([181] and [182], published in 1922 and 1924, respectively). Therefore, these relations usually are referred to as Friedmann's equations. A few years after Friedmann's publications these equations were also derived independently by G. Lemaître. Therefore, the solutions of the equations are called either "Friedmann world models" or "Friedmann-Lemaître models."

For some purposes it is convenient to rewrite Friedmann's equations in different forms. Making use of the "Hubble parameter"

$$H(t) = \frac{\dot{a}(t)}{a(t)} \tag{3.8}$$

Equation (3.5) can be written as

$$\frac{\dot{a}^2(t)}{a^2(t)} = H^2 = \frac{8\pi G}{3}\rho(t) + \frac{\Lambda c^2}{3} - \frac{kc^2}{a^2(t)} \tag{3.9}$$

Furthermore, by defining the dimensionless density parameters

$$\Omega_M = \frac{8\pi G}{3H_0^2}\rho_0,\ \Omega_\Lambda = \frac{\Lambda c^2}{3H_0^2},\ \Omega_k = -\frac{kc^2}{a_o^2 H_0^2} \tag{3.10}$$

(where the index 0 denotes the present-day values of the corresponding variables) we obtain from Eq. (3.5):

$$\Omega_M + \Omega_\Lambda + \Omega_k = 1 \tag{3.11}$$

H_0, the present-day value of the Hubble parameter, is the slope of the local redshift-distance relation (Sect. 1.2). Therefore, it is called the "Hubble constant." H_0 has the dimension s^{-1}, but usually is expressed in units of $\mathrm{km\ s^{-1}\ Mpc^{-1}}$. Instead of H_0 often the dimensionless quantity $h = H_0/(100\ \mathrm{km\ s^{-1}\ Mpc^{-1}})$ is used. Since the Hubble constant is less accurately known than other cosmic parameters, numerical values or relations which depend on the exact value of H_0 are often given with h included as a free parameter.

Modern astronomical observations (summarized, e.g., in [491]) give for the parameters defined above $\Omega_k \approx 0$, $\Omega_M \approx 0.28$, $\Omega_\Lambda \approx 0.72$, and $h \approx 0.73$. The observational mean errors of Ω_k, Ω_M, and Ω_Λ at present are assumed to be of the order 0.04. Due to possible systematic errors, the uncertainty of h is about twice as large. The parameters listed above result in a cosmic age of 13.7 Gyr. The corresponding cosmic scale as a function of cosmic age is plotted in Fig. 3.1. Table 3.1 lists, for orientation, the look-back time (time since the emission of the observed light) and the cosmic age as a function of redshift for these parameters.

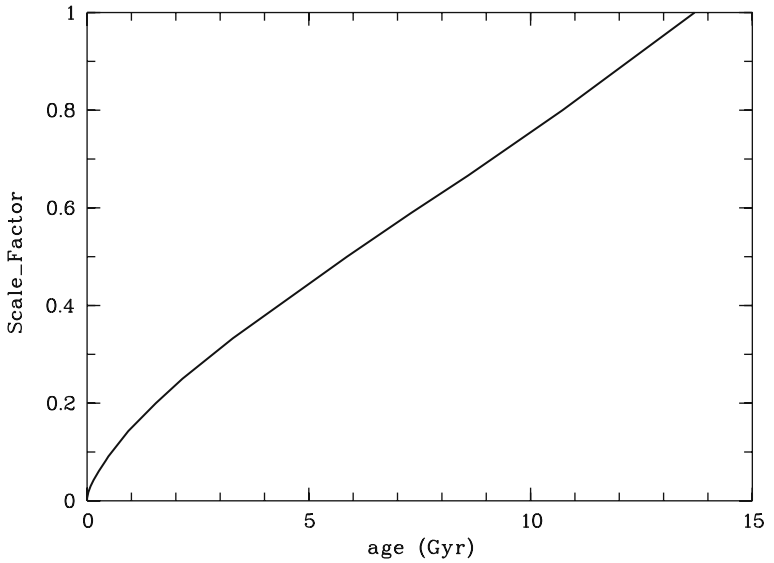

Fig. 3.1 The normalized cosmic scale factor (a/a_0) as a function of the cosmic age for a Friedmann universe with $\Omega_k = 0$, $\Omega_M = 0.28$, $\Omega_\Lambda = 0.72$ and $h = 0.73$

3.1.2 Properties of the Friedmann-Lemaître Models

In this section various properties of the Friedmann world models are discussed. These properties can be deduced without actually solving the equations.

From Eqs. (3.6) and (3.5) it is evident that (if $\rho \neq 0$) a static solution (i.e., $\dot{a} = 0, \ddot{a} = 0$) is possible only for $k = +1$, and only for special values of Λ and ρ. The corresponding static solution is Einstein's static universe. The observed cosmic expansion and the observed present-day mass-energy density ρ rule out this static solution. The observed cosmic density also rules out all solutions of Friedmann's equations which are not expanding from $a = 0$. Thus, among the non-static, expanding solutions of Friedmann's equations, only the "big bang" models are consistent with the current astronomical data.

During the expansion, ρ decreases proportional to a^{-3} or (if radiation or relativistic particles dominate the density) even faster. Therefore, at the beginning of the expansion the first term in Eq. (3.5) is always dominant. Moreover, since $\rho \to \infty$ for $a \to 0$, big-bang models (with $\rho \neq 0$) start the expansion with an infinite expansion rate. For $\rho + 3pc^{-2} > 0$ initially the expansion is always decelerated ($\ddot{a} < 0$). While the cosmos is expanding the density term and the curvature term in Eq. (3.5) both decrease, while (for a constant Λ) the Λ term remains constant. Therefore, if $\Lambda \neq 0$, big-bang models eventually reach a stage where Λ determines the expansion rate, while all other terms become negligible.

An important property of the big-bang models is the finite age of the universe. This also means that light can reach us only from those regions of the cosmos where

the light traveling time does not exceed the cosmic age. The distance R_h to the edge of the volume from which light can have reached us (and within which a past causal connection has been possible) is called the "particle horizon" (or simply the "horizon"). Obviously, we have no information on the part of the universe beyond the particle horizon. In the Friedmann–Lemaître big-bang models, the horizon distance was zero at the beginning of the expansion, and has been growing ever since.

Table 3.1 Look-back time and cosmic age as a function of redshift for a Friedmann universe with $\Omega_k = 0$, $\Omega_M = 0.28$, $\Omega_\Lambda = 0.72$, and $h = 0.73$

Redshift	0.5	1.0	2.0	3.0	4.0	6.0	10	20
Look-back time (Gyr)	5.11	7.84	10.41	11.55	12.15	12.77	13.22	13.52
Cosmic age (Gyr)	8.59	5.86	3.29	2.15	1.55	0.93	0.48	0.18

The Friedmann equations also show that for $\Lambda < 0$ (and if $\Lambda = 0$, for $k = +1$) the expansion rate $\dot{a}$ has to change its sign at some epoch. Thus, for big-bang models with such parameter values the scale factor a reaches a maximum, before the expansion is reversed and the cosmos starts to collapse again to $a = 0$. If $\Lambda > 0$ (and $\rho + 3pc^{-2} > 0$) at some epoch $\ddot{a}$ changes its sign and the expansion eventually becomes accelerated. For k = 0 the change of sign of $\ddot{a}$ occurs at the redshift $z = (2\Omega_\Lambda/\Omega_M)^{1/3} - 1$. Thus, with the cosmic parameters quoted above, for our actual cosmos $\ddot{a}$ changed sign at a redshift of $z \approx 0.73$ or about 6 billion years ago.

In order to derive solutions of Friedmann's equations, the equation of state of the cosmic matter $p = p(\rho)$ (characterized by the quantity $w = p/\rho c^2$) has to be known. Since today the velocities of most particles and masses are much smaller than the velocity of light, and since the dark matter exerts no pressure, we have $p \ll \rho c^2$, and $p = 0$ and $w = 0$ are good approximations for the matter of our present cosmos. However, as pointed out above, the cosmic mass-energy of the present-day cosmos is dominated by the dark energy. For dark energy due to a cosmological constant, we have $p = -\rho c^2$ and $w = -1$. For quintessence dark energy (Sect. 2.6) we have $-1 < w < 0$. w also does not vanish if radiation is important. For a black body radiation field we have $\rho c^2 = 3p = \mathrm{a}_R T^4$, where a_R is the radiation constant. Thus, for radiation we have $w = 1/3$.

During the cosmic expansion the wavelength λ of the radiation increases $\sim a$. Since according to Wien's law $T \propto \lambda^{-1}$ the temperature of a black-body radiation field decreases $\propto a^{-1}$ and the energy density of radiation decreases $\propto a^{-4}$. Thus, during the cosmic expansion the energy density of the radiation decreases faster than that of the matter (which changes $\propto a^{-3}$). Obviously, at sufficiently early cosmic epochs the energy density of the radiation field must have exceeded that of the matter. Matter consisting of relativistic particles (i.e., particles having a mass $\gg$ than their rest mass) behaves like the radiation field. But, among the common elementary particles with relativistic velocities only the neutrinos have a sufficiently high density to contribute significantly to the energy density of the early universe. Since the neutrino rest masses are still somewhat uncertain, the neutrino contribution is not well known (Sect. 2.7). But in spite of these uncertainties (and with the current limits of the neutrino rest masses) it can be shown (e.g., [319]) that

the neutrino contribution to the energy density of the early universe has always been smaller than the contribution of the photons.

3.1.3 Some Analytic Solutions

Solutions of Friedmann's equations are called "world models." Because of the different components of the cosmic mass-energy density (with different equations of state), solving the Friedmann equations normally requires numeric integrations. However, for various simple cases there exist analytic solutions, which are useful for estimating and understanding the qualitative behavior of cosmological models. The following are among these analytic solutions.

The "empty cosmos"

If we assume $\Lambda = 0$, $k = -1$, and $\Omega_M = 0$, Eq. (3.5) becomes $\dot{a}^2(t) = c^2$, which has the solution

$$a(t) = a_0 t / t_0 \tag{3.12}$$

describing an expansion with a constant rate ($\ddot{a} \equiv 0$). Our existence—and other observational facts—obviously rule out an empty cosmos. However, the empty universe is an interesting reference model, since (as pointed out above) about 6 billion years ago our actual cosmos went through a slow change of the sign of $\ddot{a}$. Thus, during a significant fraction of our cosmic history our universe was behaving approximately like the "empty cosmos."

The Einstein–De Sitter solution

This case assumes $\Lambda = 0, k = 0, \Omega_M = 1, w = p/(\rho c^2) = 0$. The corresponding solution

$$a(t) = a_0 (t/t_0)^{2/3} = a_0 (6\pi G \rho_0)^{1/3} t^{2/3} \tag{3.13}$$

approximates the behavior of our cosmos during phases where the energy density of the (dark) matter already exceeded that of the radiation, while the Λ term was still unimportant.

De Sitter's inflationary solution

If we assume $\Lambda > 0, k = 0$, and $\Omega_M = 0$, we get the solution:

$$a(t) = a_0 \exp\left(\frac{t}{3^{1/2} \Lambda^{-1/2} c^{-1}}\right) \tag{3.14}$$

On first glance, this case again appears unrealistic since a vanishing matter density (or at least $\Omega_M \ll \Omega_\Lambda$) is assumed, which is not (yet) the case at present. However, as will be pointed out in Sect. 3.1.4, there are reasons to assume that during an earlier cosmic epoch (when the universe was between about 10^{-35} and 10^{-32} seconds old) Ω_Λ temporarily became much larger than Ω_M. During that "inflationary" period our cosmos probably expanded exponentially according to Eq. (3.14).

The "radiation universe"

As noted above, in the very early universe the mass–energy density of the radiation field (and of the relativistic neutrinos) provided the dominant term in Eq. (3.5). For that epoch $\Lambda = 0, k = 0, \Omega_M = 1$, and $w = 1/3$ is a good approximation. The resulting solution of Friedmann's equations is:

$$a(t) = a_0(t/t_0)^{1/2} = a_0 \left(\frac{32\pi G\rho_0}{3} \right)^{1/4} t^{1/2} \tag{3.15}$$

3.1.3.1 The horizon distance of the analytical solutions

By inserting the analytic solutions $a(t)$ listed above into Eq. (3.4) and by integrating from $t = 0$ to a given time t, it is possible to derive simple analytic expressions for the horizon R_h as a function of the time t. In this way we obviously get (e.g.) for the Einstein–De Sitter solution $R_h = 3ct$, and for the radiation universe $R_h = 2ct$.

3.1.4 The Cosmic Expansion at Different Epochs

The analytic solutions of Friedmann's equations listed above can be used to characterize and to understand several distinct periods of the cosmic expansion history. As pointed out in Sects. 3.1.1 and 3.1.2 (due to the $\propto (1 + z)^4$ redshift dependence of the radiation energy density) immediately after the big bang the cosmic mass-energy density was dominated by the electromagnetic radiation field. Therefore, at the beginning the expansion is described well by the expansion law of the radiation universe given by Eq. (3.15).

However, with decreasing redshift the energy density of the photons (and of the relativistic neutrinos) decreases $\propto (1 + z)^4$, while the matter density decreases $\propto (1 + z)^3$. As a result, the ratio between energy densities of the radiation and of the matter decreases $\propto (1 + z)$. Therefore, at some point in time the matter density became larger than the mass-energy density of the radiation (plus neutrinos). From the present matter and radiation densities and the above relations (and assuming $\Omega_M = 0.3$) it follows that this must have happened at a redshift of $z \approx$ or 3500 about 4.7×10^4 years after the big bang. After the epoch of "matter-radiation equality" the expansion rate was determined by the (dark) matter density, although the pressure of the cosmic plasma initially still was due to the radiation field. For this stage the expansion rate given by Eq. (3.13) is a good approximation. Finally, about 6 billion years ago, Ω_M dropped below Ω_Λ and the dark energy started to dominate the mass-energy density, $\ddot{a}$ changed its sign, $\dot{a}$ reached a minimum value, and the present phase of accelerated expansion begun.

Various basic properties of the present-day universe, such as its large-scale homogeneity, its flatness, and its relatively high specific entropy, can be readily explained if we assume that the universe went through a temporary phase of exponential

expansion, as described by the de Sitter solution of the Friedmann equations (3.14). Therefore, it is generally assumed that approximately between 10^{-35} s and 10^{-32} s after the big bang the radiation-dominated period was interrupted by an "inflationary" phase, where the vacuum energy was dominant, and where an exponential expansion according to Eq. (3.14) took place. The total relative scale increase of the universe during inflation has been estimated (from the present entropy content of the universe) to have been of the order e^{100} or $\approx 10^{43}$ (e.g., Chap. 9 of [60]). Like all other scales, the scale corresponding to the horizon which had been reached by the beginning of the inflationary period was enormously expanded. In fact, the distance corresponding to the horizon just before the onset of inflation became much larger than the present-day horizon distance. Therefore, if the assumption of the inflationary phase is correct, in the very early universe regions much larger than the present horizon were causally connected. This may explain the large-scale homogeneity of the present universe. Moreover, if our space geometry is not flat, during inflation the radius of curvature was expanded to a value that is much larger than the present horizon. Even if the radius of curvature of our space is not infinite, if inflation took place, the present radius of curvature is so much larger than the horizon that the curvature is negligible (and unmeasurable) on observable scales.

The physical mechanism for the temporary increase of the vacuum energy density is not clear at present. It could have been a loss of equilibrium during a phase transition of the vacuum or a (so far unknown) physical field. Since the physical cause of inflation is not understood, the exact details of this cosmic event remain unclear. But there are no plausible alternatives for explaining the observed effects noted above. Therefore, inflation has become a basic ingredient of practically all current cosmological models.

During the first minute of the cosmic evolution the (visible) cosmic matter consisted of free elementary particles in thermal equilibrium. Atoms did not yet exist and the number ratio between protons and neutrons was determined by the temperature. During the following few minutes the temperature dropped below the value needed for maintaining the neutron–proton equilibrium, and the free neutrons started to decay. At the same time neutrons and protons started to form atomic nuclei. This "primordial" or "big-bang" nucleosynthesis resulted mainly in the formation of ^{4}He nuclei and very small amounts of ^{2}H (deuterium), ^{3}He, and ^{7}Li. Apart from ^{4}He, these light nuclei are not very stable. Under equilibrium conditions they form ^{4}He through collisional nuclear reactions. Only the rapid density decrease during that epoch made it possible that measurable quantities of these light nuclei survived. Therefore, the amount of light elements produced by the primordial nucleosynthesis depends critically on the baryon density and the cosmic expansion rate at this epoch. In a lower density plasma there are fewer collisions. This increases the chances of the survival of ^{2}H, and ^{3}He nuclei. Since the relevant nuclear reaction rates are reasonably well known, the abundance (relative to hydrogen) of the light nuclei can be calculated as a function of the baryon density. Moreover, the present-day abundance of these nuclei can be determined from observations of the interstellar and intergalactic medium (^{2}H, ^{3}He) or in stellar atmospheres (^{7}Li)

(e.g., [80], [101])[1]. A comparison of the observed abundances with the big-bang prediction can be used to determine the baryon density at a redshift $z \approx 10^9$, i.e., for an epoch which cannot be directly observed, since at that time the universe was still (very) opaque.

Another important milestone in the cosmic evolution was the epoch when the universe became transparent to most of the electromagnetic radiation. This occurred at a redshift $z \approx 1089$ ([47]). Since this event depended more on the (very accurately known) temperature history of the universe than on the matter density, its redshift is much more precisely known than that of the epoch of the matter-radiation equality. But there is no doubt that the transition to a transparent universe occurred when the cosmic mass-energy density was already dominated by the matter, while the contribution of the dark energy (with $w = -1$) was still negligible. Before that epoch the two essential components of the matter, H and He, were practically fully ionized. Because of the large number of photons in the universe ($n_{photons} \approx 1.6 \times 10^9 \, n_{baryons}$) the ionization could be maintained by the Wien wing of the black-body spectrum. Therefore, the universe remained ionized until the temperatures had dropped to a few thousand degrees. But then recombination occurred very rapidly, and the universe became transparent for most photons of the prevailing radiation field during a very short time interval. The transition period corresponded approximately to the redshift interval $1200 < z < 1000$. As soon as the universe had become transparent, the baryonic matter and the radiation field decoupled. Therefore, this period is called either the "epoch of decoupling" or the "epoch of recombination." Since the cosmic energy density was not affected by the recombination, this event had practically no effect on the overall dynamics of the expanding cosmos. However, it had important consequences for the formation of structure (see Sect. 3.2).

As pointed out in Sect. 2.4, today most of the intergalactic medium is ionized again. This is not surprising since present-day's hot stars and AGN produce copious amounts of UV photons which can ionize the IGM. Moreover, compared to the epoch corresponding to $z \approx 1100$, the density of the intergalactic gas is much lower and the electron temperature is much higher. Therefore, today the recombination rate is very low. As discussed in detail in Chap. 9, the reionization of the IGM probably started as soon as the first stars were formed.

3.1.5 Astronomy in the Expanding Universe

Apart from the redshift of the observed spectra, the expansion of the universe has little consequences for observations of the local universe and of nearby galaxies

[1] While the measurements for ^{2}H and ^{3}He, although technically difficult, are relatively straightforward, the derivation of the primordial ^{7}Li abundance is more problematic because the observed abundance may be affected by poorly understood diffusion effects in the stellar interiors (e.g., [277]).

and galaxy clusters with $z \ll 1$. However, when observing distant galaxies, the expansion has a profound influence on the observations. A useful tool for calculating the effects of the expansion on astronomical observations is the "expansion function" $E(z) = H(z)/H_0$, which describes the expansion as a function of the observable redshift (instead of time). Since $E(z)$ is the Hubble parameter divided by the Hubble constant, we get from Eq. (3.8)

$$\dot{a}(z) = \frac{\mathrm{d}a}{\mathrm{d}t} = a(z)H_0E(z). \tag{3.16}$$

From this relation we get for $p = 0$, $\rho \propto a^{-3}$ (using Eqs. (3.10) and (3.9)):

$$E(z) = \left[\Omega_M(1+z)^3 + \Omega_\Lambda + \Omega_k(1+z)^2\right]^{1/2} \tag{3.17}$$

or, for a flat ($k = 0$) geometry:

$$E(z) = \left[\Omega_M(1+z)^3 + \Omega_\Lambda\right]^{1/2} \tag{3.18}$$

The derivation of Eqs. (3.17) and (3.18) assumed implicitly that Λ is a constant, and that the equation of state of the dark energy is given by $p = w\,\rho\,c^2 = -\rho\,c^2$, or $w = -1$. If $w \neq -1$, Eq. (3.18) has to be replaced by

$$E(z) = \left[\Omega_M(1+z)^3 + \Omega_\Lambda(1+z)^{3(1+w)}\right]^{1/2} \tag{3.19}$$

Equation 3.17 can be used to calculate various important quantities of practical interest to astronomers. By integrating Eq. (3.16) and making use of $a = a_0(1+z)^{-1}$ we immediately get the age of the universe as a function of the redshift:

$$t(z) = H_0^{-1}\int_z^\infty \frac{\mathrm{d}z}{(1+z)E(z)} \tag{3.20}$$

In many cases this integral can be solved analytically. For our present universe we have (at least with good approximation) $k = 0$. In this case the age of the universe as a function of redshift becomes

$$t(z) = \frac{2H_0^{-1}}{3(1-\Omega_M)^{1/2}}\sinh^{-1}\left[\left(\frac{1-\Omega_M}{\Omega_M}\right)^{1/2}(z+1)^{-3/2}\right] \tag{3.21}$$

where $\sinh^{-1}$ denotes the inverse function of the function sinh.

Another quantity of significant interest is the observed angular extent Θ of an astronomical object of a fixed linear size S. From the definition of r (3.2) we get

$$\Theta(z) = \frac{S}{a(z)r(z)} = \frac{S(1+z)}{a_0 r(z)}. \tag{3.22}$$

Equation (3.22) shows the behavior of the angular size with increasing redshift. For $z \ll 1$ the angular extent decreases linearly with the distance (as in a static space with a flat geometry). For large values of z the coordinate distance r in the denominator remains finite (as r is limited by the distance which the light can have traveled since the beginning of the expansion), while $1+z$ in the numerator approaches infinity. Therefore, at high z the angular size is increasing with z, eventually becoming about proportional to z. Thus, for any object of a given linear size there exists a minimum angular extent under which it can appear on the sky.

In order to determine the angular extent and to evaluate Eq. (3.22) qualitatively, $r(z)$ has to be calculated. This can be done by integrating dr over the distance to z. For a flat universe (where $r = r_D$), we get (using Eqs. (3.4), (3.16), and $a = a_0(1+z)^{-1}$):

$$r(z) = \int \frac{c\,\mathrm{d}t}{a} = \int \frac{c\,\mathrm{d}a}{a\,\dot{a}} = \frac{c}{H_0 a_0} \int_0^z \frac{\mathrm{d}z}{E(z)}. \tag{3.23}$$

For the cosmic parameters assumed at present ($\Omega_k = 0$, $\Omega_M = 0.28$, $\Omega_\Lambda = 0.72$ and $h = 0.73$) the expected angular extend of an object with a linear size of 10 kpc (corresponding approximately to the distance between the solar system and the center of our Milky Way galaxy) is given in Fig. 3.2. As shown by this figure, the angular extent of a galaxy with a linear size of 10 kpc is always > 1 arcsec. The minimum angular extent is observed if the object has a redshift $z \approx 1.7$. Numerical results for other cosmic parameters can be calculated from Eq. (3.22), or derived from graphs reproduced, e.g., in [405].

At first glance it may appear surprising that for very distant objects of given physical size the angular extent increases with the distance. However, this result is readily understood by the fact that the high-redshift objects were still close to us,

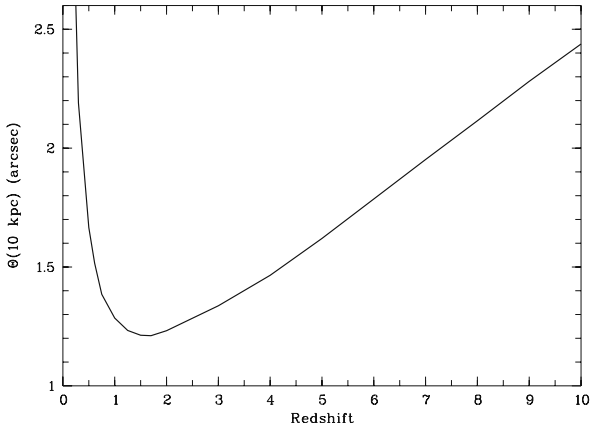

Fig. 3.2 The angular extent Θ of an object with a linear size of 10 kpc as a function of the redshift z in a Friedmann universe with $\Omega_k = 0$, $\Omega_M = 0.28$, $\Omega_\Lambda = 0.72$ and $h = 0.73$

when the observed light was emitted. Objects at intermediate redshifts emitted the observed light, after the initial very rapid expansion. At that time they were already far away from us. Hence they appear smaller. Finally, in the local universe ($z \ll 1$) the epoch difference (and consequently the distance difference) between the light emission and its observation becomes negligible. Consequently, the angular distance behaves like in a static universe.

The increase of the angular extent of astronomical objects at large redshifts has important consequences for the observations of high-redshift galaxies. As noted above, galaxies with sizes of the order 10 kpc remain always larger than 1 arcsec. This means that with modern techniques large galaxies can be spatially resolved at all distances. On the other hand, the increase of the angular size (and the angular surface) at large redshifts also means that the surface brightness drops rapidly with redshift. For nearby galaxies ($z \ll 1$) with a given luminosity the observed total flux (power per m^{-2}) with good approximation drops with the distance proportional to r^{-2}. However, since for $z \ll 1$ the angular extent decreases $\propto r^{-1}$, the solid angle, under which a nearby object is observed, also decreases like r^{-2}. As a result, the surface brightness on the sky of nearby galaxies (integrated over the whole spectrum) remains essentially constant.

To calculate the surface brightness of distant galaxies, we first have to derive the correct relation for the decrease of the flux with distance in an expanding universe. If the object emits isotropically N photons, the number of photons (eventually) received by us per unit surface obviously is

$$N_R = \frac{N}{4\pi a_0^2 r^2(z)}. \tag{3.24}$$

Moreover, because of the cosmic expansion, the rate at which the photons are received is reduced (relative to a static universe) by the factor $(1+z)^{-1}$. In addition, the redshift reduces the energy $h\nu$ of the individual photons by the same factor $(1+z)^{-1}$. Thus, the total energy received per unit surface (i.e., the observed radiation flux) becomes

$$F = \frac{L}{4\pi a_0^2 r^2(z)(1+z)^2}. \tag{3.25}$$

As shown by Eq. (3.25), in an expanding universe the flux from galaxies drops proportional to $r^{-2}(1+z)^{-2}$, while (according to Eq. 3.22) their surface changes proportional to $r^{-2}(1+z)^2$. As a result, the surface brightness of galaxies decreases $\propto (1+z)^{-4}$. This confirms that for $z \ll 1$ the surface brightness is essentially constant, while for large z the surface brightness decreases $\propto z^{-4}$. Since the surface brightness of nearby galaxies typically is similar to that of the night sky, the rapid decrease of the surface brightness with high z severely complicates the observations of high-redshift galaxies.

Another useful relation for observers of the distant expanding universe is the "proper displacement function" defined as ds/dz for $dr = d\theta = d\phi = 0$. Using

Eq. (3.16) this function can be written as:

$$\frac{\mathrm{d}s}{\mathrm{d}z} = -c\frac{\mathrm{d}t}{\mathrm{d}z} = -c\frac{\mathrm{d}a}{\mathrm{d}z}\dot{a}^{-1} = \frac{c\,a}{(1+z)\dot{a}} = \frac{c}{H_0(1+z)E(z)} \tag{3.26}$$

An important application of Eq. (3.26) is the derivation of the proper distance which corresponds to a redshift difference. Moreover, Eq. (3.26) can be used to relate surface densities on the sky to local space densities.

If, e.g., objects with a cross section σ have at the redshift z a space density of $n = n_0(1+z)^3$, the probability per unit redshift interval that a random line of sight intersects such an object obviously is:

$$\frac{\mathrm{d}P}{\mathrm{d}z} = n\sigma\frac{\mathrm{d}s}{\mathrm{d}z} = n_0\sigma\frac{c(1+z)^2}{H_0\,E(z)} \tag{3.27}$$

An expression for the proper volume at a redshift z can be obtained in the following way. From Eq. (3.26) we readily get an expression for the geometric depth δs corresponding to a redshift interval δz. The area observed within an solid angle $\delta\Omega$ at an angular size distance r obviously is $a^2r^2\delta\Omega$. Multiplying the area with $\delta s = \frac{\mathrm{d}s}{\mathrm{d}z}\delta z$, we get for the corresponding proper volume δV:

$$\delta V = \frac{ca_0^2r^2\delta\Omega\,\delta z}{H_0(1+z)^3E(z)} \tag{3.28}$$

Consequently, if the number of objects inside a comoving volume is conserved (i.e., if $n = n_0(a_0/a)^3$), and if all objects can be observed at all redshifts, the number of objects observed per steradian and per unit redshift interval is

$$\frac{\mathrm{d}N}{\mathrm{d}z} = \frac{cn_0a_0^2r^2}{H_0E(z)}. \tag{3.29}$$

Graphic presentations of these relations can be found, e.g., in [405].

3.2 The Formation and Evolution of Structure

3.2.1 The Origin of Structure

The homogeneous models of our universe described in Sect. 3.1 are useful for describing the overall properties and the global evolution of the universe. However, our present universe shows structure with large local density variations, and the quantum nature of matter excludes a completely homogeneous cosmos. Therefore, the homogeneous models obviously cannot provide a complete description of our actual cosmos.

From quantum theory we expect that there are fluctuations (in space and time) of the cosmic mass-energy density. Such fluctuations of the mass-energy density result in variations of the curvature of space, which act on the matter as gravitation. Since gravitation attracts matter to regions of higher than average density, gravitation tends to amplify density fluctuations. Therefore, it is generally assumed that all cosmic structure, which we observe today originated from quantum fluctuations in the early universe, which were amplified by gravitation. Although quantum fluctuations occurred (and were amplified) from the beginning of the universe, during the inflationary epoch all previous density fluctuations were wiped out by adiabatic cooling and by the enormous scale expansion. Therefore, the present-day cosmic structure is assumed to have its origin in quantum fluctuations which occurred at the end of inflation, when the universe was reheated by the decay of the field which caused the rapid expansion. Our lack of understanding of the physics of inflation and the absence of a generally accepted theory of quantum gravity prevents a reliable theoretical prediction of the amplitude and the power spectrum of the initial fluctuations. Nevertheless, it can be shown that with plausible assumptions inflation can produce adiabatic density fluctuations with an amplitude and a power spectrum compatible with the observational data (e.g., [69]). Therefore, the current theories of cosmic structure formation are basically attempts to model the evolution of density fluctuations between the epoch of cosmic inflation and now. As initial conditions the computations normally assume a power spectrum $P(k)$ (i.e., the square of the amplitude of the fluctuations as a function of the spatial wave number) of the initial fluctuations of the form

$$P(k) \propto k^n \tag{3.30}$$

where k is the spatial wave number. The observations of the fluctuations of the cosmic microwave background are consistent with a spectrum according to Eq. (3.30) with $n \approx 1$. A power spectrum with $n = 1$ (which means $P(k) \propto k \propto \lambda^{-1}$) had been suggested independently by E. Harrison and Y. Zeldovich long before observational evidence on the shape of the spectrum could be derived. Therefore, a spectrum of the form $P(k) \propto k$ is called a "Harrison-Zeldovich" spectrum. For $n = 1$ we get $P \to 0$ for $k \to 0$. Thus, for a Harrison-Zeldovich spectrum the universe becomes homogeneous at large scales. The agreement of this prediction with the observations was one of the reasons why a power-law spectrum according to Eq. (3.30) with $n = 1$ was originally suggested. The observations of the cosmic microwave background and other data exclude an index n which is much different from unity. On the other hand, some theoretical predictions suggest an index $n \neq 1$. Therefore, deriving the exact value of n provides a valuable test of current cosmological theories.

In order to fully describe the fluctuations, in addition to the power spectrum, the amplitude distribution of the fluctuations at a given wavenumber has to be specified. The observations of the CMB indicate that a Gaussian amplitude distribution is at least a good approximation. Therefore, normally a Gaussian distribution is assumed and the amplitude is normalized to reproduce the observed amplitude of the cosmic microwave background fluctuations.

Although calculations have been carried out for adiabatic and non-adiabatic fluctuations, CMB observations show that the cosmic structure formation was dominated by adiabatic fluctuations. Therefore, only the evolution of adiabatic fluctuations will be discussed here.

The main results of the model computations of the evolution of cosmic structure from these initial conditions are briefly summarized in the next section. A more detailed discussion of the theory and the techniques used for modeling of the early phases of structure formation can be found, e.g., in [335] and [369], while a comprehensive description of the later stages is given in [319].

3.2.2 Modeling the Evolution of Structure

3.2.2.1 Methods and Strategies

Before the formation of galaxies the amplitudes of the large-scale cosmic density fluctuations were small. Therefore, at first glance Newtonian physics appears to be a sufficient approximation for treating the gravitational effects. However, at the end of the inflation the (new) comoving horizon was much smaller than the sizes of volumes which eventually formed galaxies and galaxy clusters. Using the relations and numbers discussed in Sect. 3.1, one finds that the comoving size of the new horizon at the end of inflation was just a few meters. Even if present structures originated much later, it is clear that galaxies have formed from density fluctuations which initially were very much larger than the horizon. Under these conditions Newtonian physics cannot be used to model the evolution of the fluctuations. At least for the early evolutionary phases Einstein's equations (Eq. 3.1) have to be applied. However, since the evolution of spatial variations of the density are to be studied, the assumption of homogeneity, used Sect. 3.1, has to be dropped. If density fluctuations are present, the assumption of a universal time is no longer strictly valid and (since the curvature varies in space) the components of the metric tensor become functions of the space coordinates. On the other hand, from the observations of the cosmic microwave background we know that during the early phases of cosmic evolution the amplitudes of density variations were small. Therefore, it seems plausible to assume that the deviations between Friedmann models and the real early cosmos can be treated as small perturbations. In this case, the geometry of the early structured universe can be approximated by small, linear deviations from the Robertson–Walker metric defined by Eq. (3.2). Since during the early cosmic epochs immediately after inflation the curvature term in the Friedmann equation certainly was negligible, a flat mean geometry can be assumed and Cartesian coordinates can be used for the unperturbed 3-dimensional space. These assumptions lead to an approximate metric tensor and a corresponding line element of the form

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\sum(\delta_{ik} + h_{ik})\mathrm{d}x^i\,\mathrm{d}x^k \tag{3.31}$$

where i, k =1,3, $\delta_{ik} = 1$ if $i = k$, $\delta_{ik} = 0$ if $i \neq k$, and $h_{ik} \ll 1$. The metric tensor components h_{ik} in Eq. (3.31) are functions of space and time. Furthermore, it is usually assumed that the functions h_{ik} can be decomposed into Fourier components characterized by wave numbers $k = 2\pi/\lambda$, where λ is the spatial wavelength. Because of their orthogonality the individual Fourier components then can be treated independently. As described, e.g., in [575], with these assumptions Einstein's equations result in a set of differential equations where the dependent variables are the functions h_{ik}, the local density variations, and the local matter velocity. With the further assumption that the cosmic energy-density is due to radiation, dark matter, and baryonic matter (which is assumed to be coupled to the radiation field by electron scattering), one eventually obtains a set of differential equations for $\delta\rho/\bar{\rho}$, where $\delta\rho = \rho - \bar{\rho}$ is the local deviation of the density ρ from the mean density $\bar{\rho}$. These equations (e.g., Eq. 75 of [575]) describe damped and forced harmonic oscillations of the cosmic matter density. A close look at the different terms shows that the damping is due to the cosmic expansion and radiation effects, while the forcing term is caused by the variable (pseudo-Newtonian) gravitational potential of the dark matter density distribution. Depending on the local physical conditions and the cosmic epoch, the solutions of these equations can be either a growth of $\delta\rho/\bar{\rho}$, or oscillations of $\delta\rho/\bar{\rho}$ with time. Due to the small amplitude of the variations and since the physics of the early cosmic plasma (consisting of hydrogen and helium only) is well understood, numerical solutions of the equations describing the linear evolution can be obtained straightforward. Therefore, computations of the early structure evolution are rather reliable and accurate, and less complex than many other model computations in astrophysics. Suitable computer codes are publicly available (see, e.g., [576]).

The linear theory outlined above obviously is valid only as long as $\delta\rho/\bar{\rho} \ll 1$. In today's universe the density contrast on galactic scales is larger. Therefore, the linear approach cannot be used to follow the evolution of structure up to the present time, and other methods have to be used when $\delta\rho/\bar{\rho}$ approaches unity. Fortunately, this happens long after all fluctuations of astronomical relevance have reached their horizon and, in fact, long after the epoch of decoupling of baryonic matter and radiation. Thus, Newtonian physics can be used to model the final phases of the evolution of structure. The Newtonian simulations are normally started at a redshift well after the epoch of decoupling (typically at an epoch corresponding to $z \approx 100$). In this case the initial conditions can be determined reliably from the CMB observations. A proper approach to model the late stages of structure evolution would obviously be to describe the baryonic matter using the basic equations of Newtonian hydrodynamics and to describe the dark matter as a gas consisting of collisionless, non-interacting particles, with both components being coupled by their joint gravitational field. However, although hydrodynamic computations of this evolutionary phase have been carried out, an adequate treatment of the large range of scales involved is difficult with hydrodynamic numerical codes. Moreover, since the matter density is dominated by the dark matter, on sufficiently large scales the density evolution is determined essentially by the dark matter alone. Therefore, the usual approach for these stages is to calculate the dynamic evolution of the

dark matter alone and to add on the evolution of the baryonic matter using various approximations.

3.2.2.2 Fluctuations in the Opaque Universe

This section summarizes the basic results of the numerical modeling for the early evolutionary phases. Qualitatively, the numerical results can be understood from simple physical arguments. As an example we will follow the behavior of a density fluctuation $\delta\rho/\bar{\rho} > 0$ ($\delta\rho \ll \bar{\rho}$) with a total mass of the order typical of a galaxy or a small galaxy cluster. In this subsection describes the history of the fluctuation from the formation to the epoch when the cosmos becomes transparent. The later evolution will be discussed in the following subsection.

As noted above, at the time of formation the size of a density fluctuation corresponding to a galactic mass was very much larger than the horizon. Therefore, initially the evolution of the density fluctuation was governed by the geometry. Local physics initially had little influence on the fate of the fluctuation. From GR (and from Eq. 3.5) it is clear that any $\delta\rho > 0$ caused at the position of the fluctuation a positive space curvature, which resulted in a locally slower expansion rate $\dot{a}$. Therefore, while the universe was expanding, at the position of a fluctuation $\delta\rho > 0$ the density was decreasing less rapidly than the average density. Correspondingly in a region with $\delta\rho < 0$ the density was decreasing faster than the average density. This means that during the epoch where the size of the fluctuation was smaller than the horizon, the amplitude of the density fluctuations was growing with time. A quantitative treatment shows that during the radiation dominated epoch $\delta\rho/\bar{\rho}$ was growing $\propto a^2 \propto (1+z)^{-2}$ (e.g., [319]). A growth according to this law is expected to last until either the horizon overtakes the size of the fluctuation, or until the dark and the visible matter start to dominate the energy density.

From the relations derived in Sect. 3.1 it can be estimated that a fluctuation with a galactic mass found itself inside the cosmic horizon well before the epoch of equality of matter and radiation energy density was reached. As soon as the fluctuation became smaller than the horizon, local Newtonian physics became important. In particular, the pressure of the photon gas now could counteract the gravitation and establish an equilibrium between the pressure forces, the forces due to gravitation and the effects of the expansion. Whether such an equilibrium was possible depended on the temperature (or the velocity of sound) and the density. Like in the case of interstellar clouds, described in Sect. 2.2.5, there exists a "Jeans mass" below which an equilibrium can exist. Because of the expanding geometry and since the pressure is essentially due to the radiation field (while the contribution of the gas particles is negligible), the expression for this Jeans mass is different from that calculated for a static geometry and an ideal gas, given by Eq. (2.12). But, taking the expansion and the radiation pressure into account, the Jeans mass of an expanding fluctuation can be readily calculated from local physics (see, e.g., Chap. 11 of [319]). It turns out that the mass of a future galaxy always fell below the Jeans mass as soon as the corresponding density fluctuation found itself inside the horizon. Therefore, for a density fluctuation corresponding to a future galaxy,

pressure equilibrium was reached and $\delta\rho/\bar{\rho}$ stopped growing as soon as the fluctuation entered the horizon. However, since up to this point in time the volume of the fluctuation had been contracting relative to the expanding mean geometry, the fluctuation overshot the equilibrium state before the contraction was stopped. As a result, $\delta\rho/\bar{\rho}$ did not simply settle to the equilibrium value, but started oscillating around this value. Because of the cosmic expansion these oscillations were damped, and their amplitude was slowly decreasing with time. Moreover, the radiation flow from compressed, hot volumes to adjacent cooler regions also resulted in the loss of pulsational energy. Because of the high opacity of the matter at this epoch, radiative damping was unimportant for large fluctuations. But oscillations of fluctuations which were not (much) larger than the mean free path of the photons were affected. In fact, since at this stage the photons provided the matter density, fluctuations which were smaller than the photon mean free path were wiped out by the escape of the photons from the overdense volumes. This radiative destruction of small fluctuations was first described by J. Silk in 1968 [476]. Therefore, this process is called "Silk damping." Numerical estimates show that in this way the fluctuations less massive than about $10^5 M_\odot$ were quenched before the radiation-dominated epoch ended.

The oscillations described above were periodic time variations of the cosmic gas density. Thus, these oscillations were closely related to the density variations caused by sound waves. Hence, these oscillations of the cosmic plasma are called "acoustic oscillations." Of course, the wavelengths of the oscillations of the cosmic plasma were much larger and their frequencies were much lower than those of audible sound waves.

Since the acoustic oscillations started when the horizon scale and the size of the fluctuation became equal, and since the horizon size depends only on the time, all oscillations of a certain wavelength started at the same cosmic time. Moreover, all oscillations of the same wavelength remained exactly in phase during the following evolution.

Once started, the acoustic oscillations could continue as long as the universe was opaque and the baryonic matter was coupled to the radiation field. Since the dark matter was not interacting directly with the radiation field, its overdensity could in principle continue to grow. However, since the mass, and thus the gravitational potential, was determined by the radiation field, $\delta\rho/\bar{\rho}$ of the dark matter grew only very weakly ($\propto \log a$) during the epoch where the energy density of radiation field was still dominant. The situation changed when the epoch of equality of radiation and matter density was reached. From then on the dark matter determined the gravitational potential. As a result, $\delta\rho/\bar{\rho}$ of the dark matter started to grow again, while the densities of the radiation and the visible matter continued to oscillate (Fig. 3.3). A simple analytic derivation shows that at this stage the amplitude of the dark-matter density fluctuations grew $\sim a \sim (1+z)^{-1}$, i.e., linearly with the scale factor (e.g., [319]). In Fig. 3.3 the different phases of the early growth of a galaxy-mass fluctuation are depicted schematically.

This linear growth $\propto a$ of the dark-matter fluctuations continued until the dark energy became dynamically dominant. Since the growth of the density fluctuations

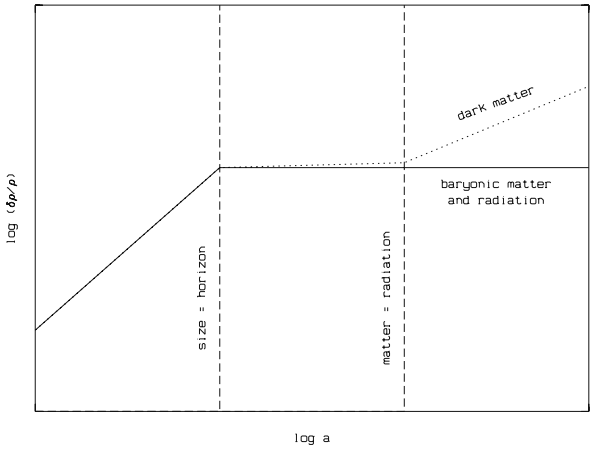

Fig. 3.3 Schematic representation of the growth of the density contrast $\delta\rho/\bar{\rho}$ of a (future galaxy-size) density fluctuation as a function of the cosmic scale $a \sim (1+z)^{-1}$ during the opaque cosmic epochs. The cosmic scale and the cosmic age increase from the left to the right. Initially, as long as the size of the fluctuation is larger than the horizon, the density contrast grows $\propto a^2$. When the fluctuation passes the horizon, it is stabilized by the radiation pressure, and $\delta\rho/\bar{\rho}$ starts to oscillate. The density contrast of the dark matter still increases somewhat. But this increase is small at this stage ($\propto \log a$) since the gravitational potential is still governed by the radiation. When the mass-energy density of the matter reaches equality with that of the radiation field, the density contrast of the dark matter starts to grow (now $\sim a$) due to its own gravitational field. The density contrast of the radiation field and of the baryonic matter continue to oscillate around an essentially constant mean value, until the cosmos becomes transparent

is an effect of the gravitational attraction of the matter, the linear growth of the fluctuations ended when the dark energy started to dominate the dynamics of the cosmos.

The dense, hot cosmic plasma was coupled to the radiation field by electron scattering. This coupling was maintained as long as the cosmic matter remained ionized. Up to this epoch the fraction of the initial fluctuation which was composed of the baryons and the radiation field continued to oscillate. However, when the dark matter decoupled from the baryons and started to collapse again, the gravitational potential of the collapsing dark matter started to influence the oscillations. The oscillating radiation field and the visible matter density experienced a gravitational dragging force of the (relative to the expanding environment) collapsing dark matter. This dragging force amplified the oscillations during the contracting phases and attenuated the oscillations during the expanding phases.

As a result of the continuing expansion the electron density was decreasing. Therefore, the mean free path of the photons was growing, and progressively larger volumes were affected by Silk damping. This led to the destruction of oscillating baryonic matter fluctuations of progressively larger sizes. On the other hand, the dark-matter fluctuations were not affected by this process. Therefore, all dark-matter fluctuations, which existed at the time of matter-radiation equality, survived and grew independently of their size.

As pointed out above, the acoustic oscillations could last until the universe became transparent and the radiation pressure disappeared. Since this was the epoch when the cosmic microwave background (CMB) was emitted, we can directly observe the acoustic oscillations of this epoch by imaging the CMB sky (see Sect. 3.2.3).

The different growth regimes outlined in Fig. 3.3 also had a profound effect on the power spectrum of the fluctuations. As explained in Sect. 3.2.1, originally the matter density fluctuations had a power spectrum (at least approximately) of the form $P(k) \propto k$. As long as the fluctuations were larger than the particle horizon, they all grew $\propto a^2$. Hence, the power spectrum was conserved during this very early epoch. However, when the horizon reached the size of a given fluctuation, the fast growth was halted. A significant growth of the amplitude of the dark-matter fluctuations continued only after the matter density started to dominate. The cosmic time (and thus the corresponding value of a) where the initial fast growth of the fluctuation amplitude was stopped, depended on the size of the fluctuation. For small fluctuations the horizon size overtook the fluctuation size earlier than in the case of large fluctuations. As a result, small fluctuations stopped growing earlier. For the power spectrum this had the effect that the amplitude and power were strongly reduced on small scales, or high values of k. Only on very large scales, where the horizon was reached after the epoch of equality of the matter and radiation density, the original $P \propto k$ power spectrum has been conserved. For smaller scales the k-dependent ending of the $\sim a^2$ growth of the fluctuation amplitude resulted in a turnover of the spectrum and a decrease of the power with increasing k, eventually approaching approximately $P(k) \propto k^{-3}$. The wave number of the maximum of the "processed" power spectrum corresponds to a comoving wave number of about 10^{-2} Mpc^{-1}. (For a quantitative treatment of the processing of the power spectrum see, e.g., [30]. For the shape of the processed power spectrum, see Fig. 3.10).

The size-dependent duration of the $\propto a^2$ growth of the density contrast obviously also influenced the amplitudes of the individual fluctuations. Interestingly, for an initial Harrison-Zeldovich power spectrum the processed amplitudes of the fluctuations became independent of their size. Thus, after passing the horizon, all fluctuations had the same amplitude. (For a simple derivation of this result see, e.g., Sects. 14.2 and 14.3 of [319].)

3.2.2.3 The Evolution After Decoupling of Matter and Radiation

After the universe had become transparent and after the baryonic matter and the radiation had decoupled, the two components evolved separately. Except for minor radiative interactions with the matter at later stages (such as the Sunyaev–Zeldovich effect) and effects of the local gravitational potential, the radiation field simply expanded, keeping its black-body spectrum. The temperature of the black-body continued (and still continues) to decrease $T \propto a^{-1}$. The baryonic matter expanded and cooled adiabatically. At the time of the decoupling the baryonic gas pressure was negligible relative to the gravitational effects on mass scales $> 10^5 M_\odot$. Therefore,

the baryonic matter of galaxy-mass fluctuations followed the dark matter in free fall and accumulated in the gravitational potential wells of the dark-matter fluctuations. (Note that the fluctuations of the dark matter had been growing significantly ($\propto a$) while the density of baryonic matter was oscillating.) At a redshift $\approx 10^2$ the baryonic matter reached a relative density enhancement similar to that of the dark matter, and $\delta\rho/\bar{\rho}$ of the dark and the baryonic matter now both increased $\propto a$. Since the maximum of the CMB spectrum soon was redshifted to IR wavelengths, and since there were no sources of visual radiation, this cosmic epoch often is referred to as the "dark age."

During the following cosmic evolution the expansion rate was decreasing, while the density contrast was increasing. As a result, those fluctuations which initially had the highest amplitudes became gravitationally bound and stopped expanding. These gravitationally bound dark-matter concentrations are called "dark-matter halos." Their properties are discussed in more detail in Sect. 3.3.1.

The particularly dense matter fluctuations eventually resulted in the formation of today's galaxies and galaxy clusters. The less dense fluctuations continued to expand. Fluctuations which had not been able to form bound objects in the past will have no chance to do so in the future, since the present accelerated expansion of the universe prevents a further linear growth of the amplitude of these fluctuations, while the density decrease due to the expansion is becoming progressively faster.

When the forces due to the expansion of the universe became negligible $\delta\rho/\bar{\rho}$ no longer increased proportional to $(1 + z)^{-1}$. Instead, the density increase became exponential (as in the case of collapsing interstellar clouds at the beginning of the star formation process). The characteristic timescale of the exponential density increase is the free-fall time t_{FF}. The exact value of the free-fall time depends on the shape of the fluctuation. But t_{ff} is always $\propto \bar{\rho}_0^{-1/2}$, where $\bar{\rho}_0$ is the initial mean density. Since in this regime a much faster density increase is possible, the assumption $\delta\rho \ll \bar{\rho}$ and the linear theory used for calculating the early stages of structure formation are no longer applicable. On the other hand, as noted already, during these late stages, Newtonian physics is a sufficient approximation (as long as the volumes studied are significantly smaller than the horizon scale). Moreover, the dynamics on large scales is determined by the dark matter, which is a pressureless fluid (equation of state $p = 0$). Therefore, models on the later "nonlinear" stages of the evolution of dark matter can be numerically simulated following the evolution of a pressureless Newtonian fluid under its own gravitation. Technically, this is achieved by assuming the dark matter to consist of a large number of point-like particles which interact only by gravitation. This technique, which was initially developed for studying the evolution of star clusters, is usually referred to as "N-body simulations." Among the most extensive calculations of this type carried out so far has been the "Millennium Simulation" [492]. This simulation has been carried out by a large international consortium. It followed the evolution of a cube of cosmic matter with a comoving size of 500 h^{-1} Mpc on a side, approximated by $\approx 10^{10}$ particles, from the redshift $z = 127$ to the present-day epoch. The resulting evolution of the dark matter structure is illustrated by Figs. 3.4 and 3.5. Figure 3.4 shows the dark-matter density

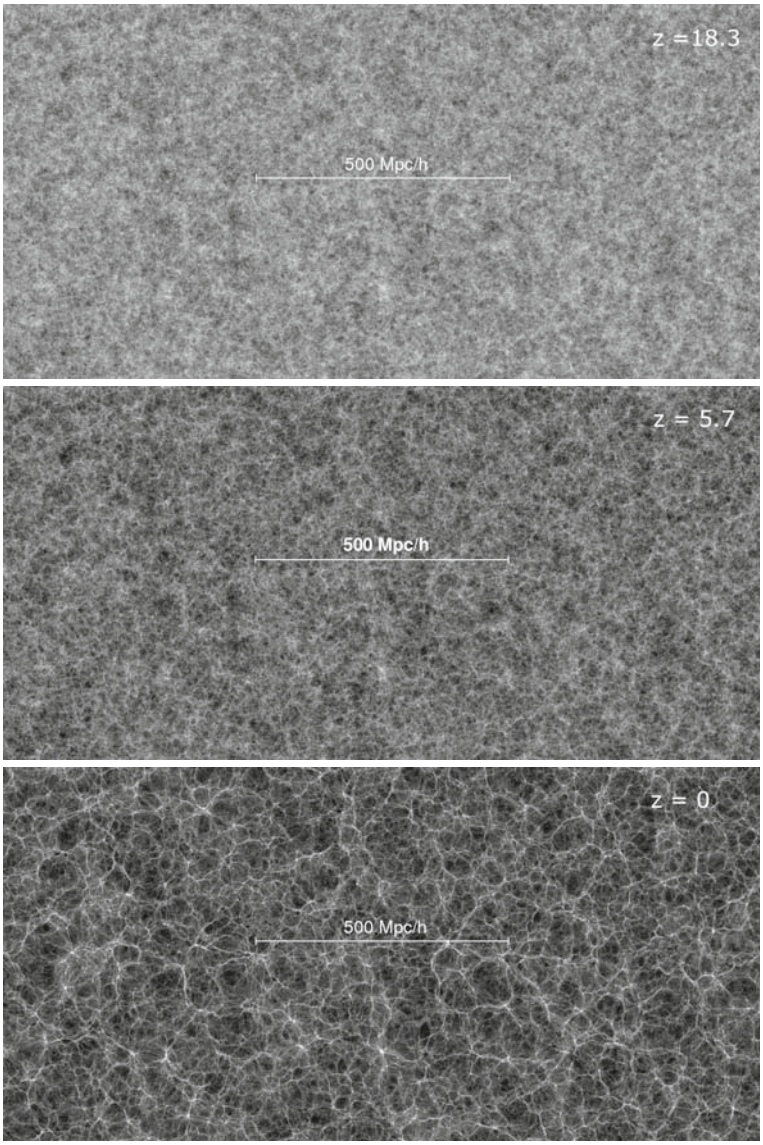

Fig. 3.4 Dark-matter density distributions derived in the Millennium Simulation [492] for different redshifts z. Bright regions correspond to higher densities. (Courtesy MPI for Astronomy, Garching, Germany)

in a (comoving) 15 Mpc thick slice of the model at the epochs corresponding to redshifts of 18.3 and 5.7, and for the present-day universe ($z = 0$). Figure 3.5 shows (for $z = 5.7$ and $z = 0$) an enlargement of the region of the prominent density peak below the center of the field reproduced in Fig. 3.4.

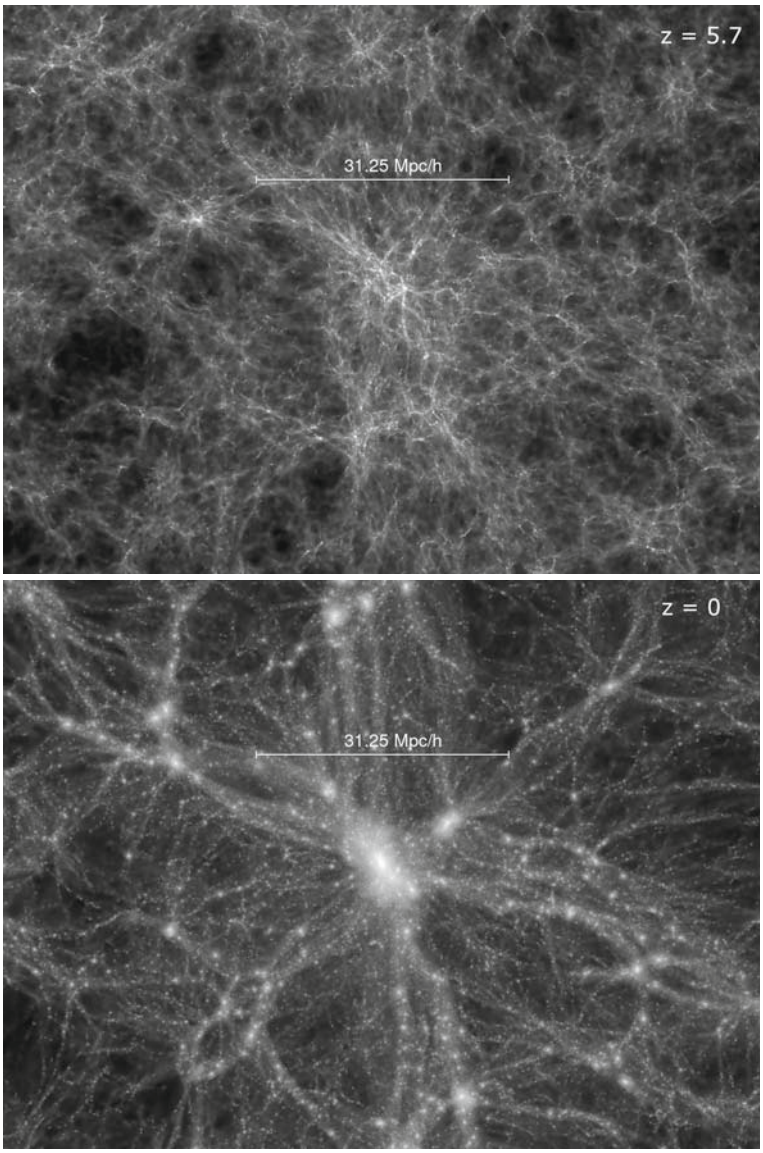

Fig. 3.5 Enlargement from Fig. 3.4. The enlargement is centered on the high-density region in the lower middle part of Fig. 3.4. (Courtesy MPI for Astronomy, Garching, Germany)

The figures show the calculated increase of the density contrast with decreasing redshift (or increasing cosmic time). They also show that the evolution of the dark-matter distribution results in a sponge-like structure observed empirically for the large-scale distribution of the galaxies (Sect. 2.3). Like the observed distribution of galaxies, the calculated dark-matter distribution shows the mass to be concentrated

in filaments and sheets. In the context of the dark-matter distribution this structure is sometimes called "the cosmic web." The highest densities are reached in the knots of the "cosmic web," where much of the total dark-matter mass is accumulating, and where the first bound systems start to form. Thus, these knots are the preferred birth places of the first galaxies and massive galaxy clusters (see Sect. 3.3).

N-body simulations can also be used to determine the characteristic density profiles of the dark-matter halos. As shown by J. F. Navarro, C. S. Frenk, and S. White [370], over most of the distance r to the center of a virialized cold dark-matter halo the density follows approximately the relation

$$\rho(r) \propto \frac{1}{(r/r_s)(1 + r/r_s)^2} \tag{3.32}$$

where r_s is a characteristic radius. Density profiles following this law are referred to as "Navarro-Frenk-White" or NFW profiles.

In order to model the evolution of the baryonic matter, the simulations of the evolution of structure of the dark matter are often combined with calculations simulating the physical processes in the baryonic matter which is accumulating in the potential wells of the dark-matter concentrations. For this purpose the most relevant processes (such as cooling, heating, radiative effects, and star formation) are approximated using analytical functions or models. This combination of N-body calculations and analytic approximations of the physics of the baryonic matter is called "semi-analytic modeling." For details on this technique see, e.g., [263, 484]. Figures 3.6 and 3.7 show results based on this method. In the case depicted in these figures the dark-matter distribution derived in the Millennium Simulation and semi-analytic methods have been used to predict the formation of galaxies and galaxy clusters. As shown by the figures, the calculated formation of galaxies follows fairly well the dark-matter distribution. (For details see [492].) This obviously explains while the sponge-like structure of the dark matter results in a qualitatively similar structure of the galaxy distribution.

3.2.3 Observational Tests

3.2.3.1 CMB Observations

As pointed out above, if the cosmic parameters and the initial conditions are known, the early stages of structure formation (up to the emission of the CMB) can be simulated rather reliably by model computations. At initial conditions usually small fluctuations with a power spectrum given by Eq. (3.30) and a Gaussian amplitude distribution are assumed. The main parameters entering the computations are Ω_M, $\Omega_{Baryons}$, Ω_Λ, Ω_k, the normalized Hubble constant h, the power spectrum index n, and a normalizing factor fixing the absolute amplitude of the fluctuations. With a given set of parameters the evolution of the fluctuations is completely determined by the physics. Therefore, for a fixed power spectrum the absolute fluctuation

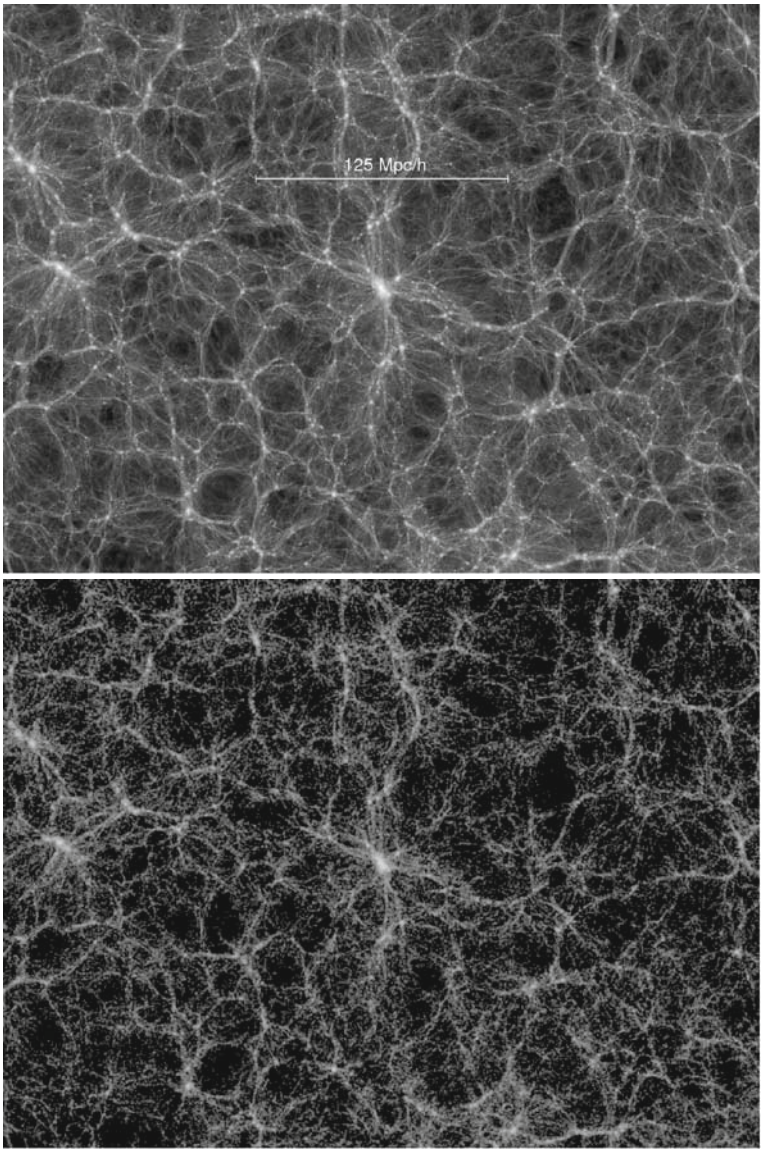

Fig. 3.6 *Upper panel*: $z = 0$ density distribution of the dark-matter found from the Millennium Simulation ([492]). *Lower panel*: Corresponding galaxy distribution resulting from semi-analytic modeling. (Courtesy MPI for Astrophysics, Garching, Germany)

amplitude can be defined by a single numerical value at an arbitrarily selected epoch and for an arbitrarily selected size. For the comparison with the observations it is of advantage to define the fluctuation amplitude at a late cosmic epoch. Therefore, the absolute fluctuation amplitude is often defined by the variance of the matter density

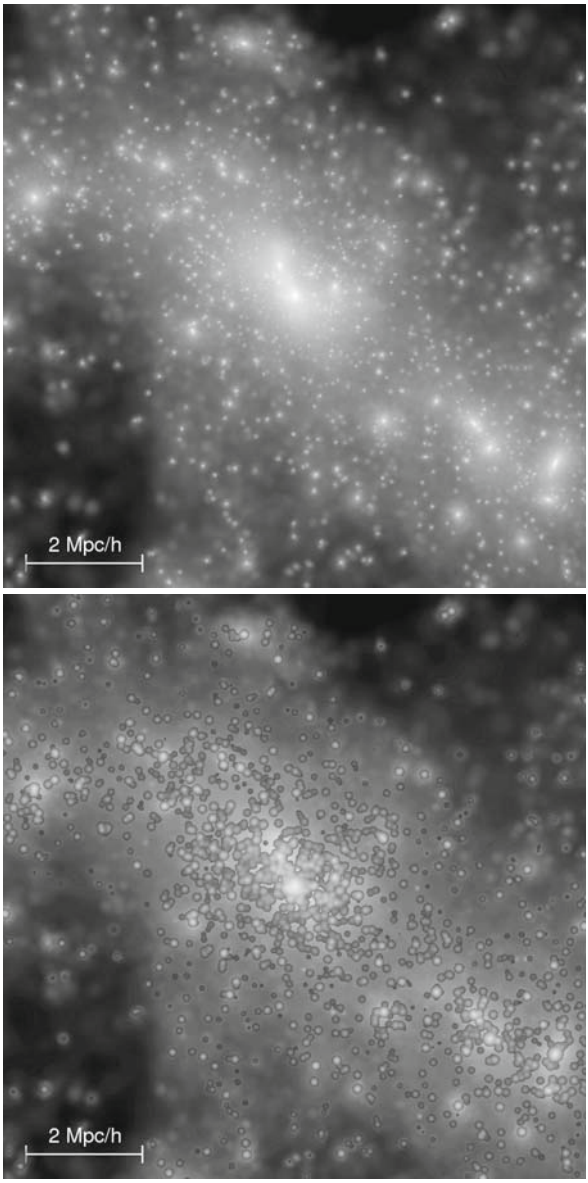

Fig. 3.7 Enlargement of the central region of Fig. 3.7. The *upper panel* again shows the dark-matter distribution. The *lower panel* shows the distribution of the galaxies (indicated by small *circles*) overlayed on the dark-matter distribution. (Courtesy MPI for Astrophysics, Garching, Germany)

contrast found in spheres of radius $8h^{-1}$ Mpc, extrapolated to the present epoch. This quantity is called σ_8. Thus,

$$\sigma_8^2 = \left\langle (\delta\rho/\bar{\rho})^2 \right\rangle_{8h^{-1}Mpc} \tag{3.33}$$

The scale $8h^{-1}$ Mpc has been chosen since the variance of the present relative number density of galaxies $\delta n_g/\bar{n}_g$ is known to be about unity at this scale. The observed numerical value of σ_8 is about 0.8 (e.g., [491]).

The model computations can be used to predict the appearance and the properties of the CMB in great detail (e.g., [369, 575]). Moreover, by performing the computations with different parameters and by comparing the results with the observations, the calculations can be used to constrain the input parameters listed above.

Among the most important empirical tests of the structure formation theory is a comparison of the predicted and the observed power spectrum of the brightness fluctuations of the CMB sky. Such comparisons are presented in Figs. 3.8 and 3.9. The predicted spectrum was calculated assuming for the fluctuations an original power

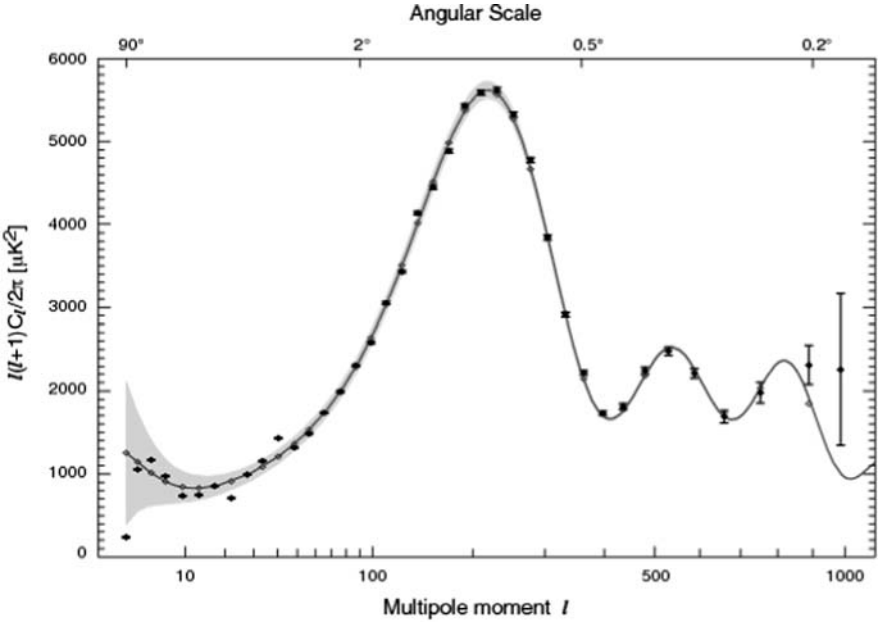

Fig. 3.8 Observed angular power spectrum of the intensity variations of the cosmic microwave background radiation. The *solid line* is the predicted power spectrum based on the best-fit cosmological parameters. The data points are the measurements of the WMAP satellite during the first 3 years of operation ([231, 491]). The *shaded area* indicates the statistical uncertainty of the observations resulting from the finite numbers of observed fluctuations. While this statistical uncertainty due to "cosmic variance" is negligible for small scales (where the data points are based on large averages) the comic variance introduces significant uncertainties at the low-frequency end, where the celestial sphere covers only a few fluctuations. (Credit: NASA/WMAP Science Team)

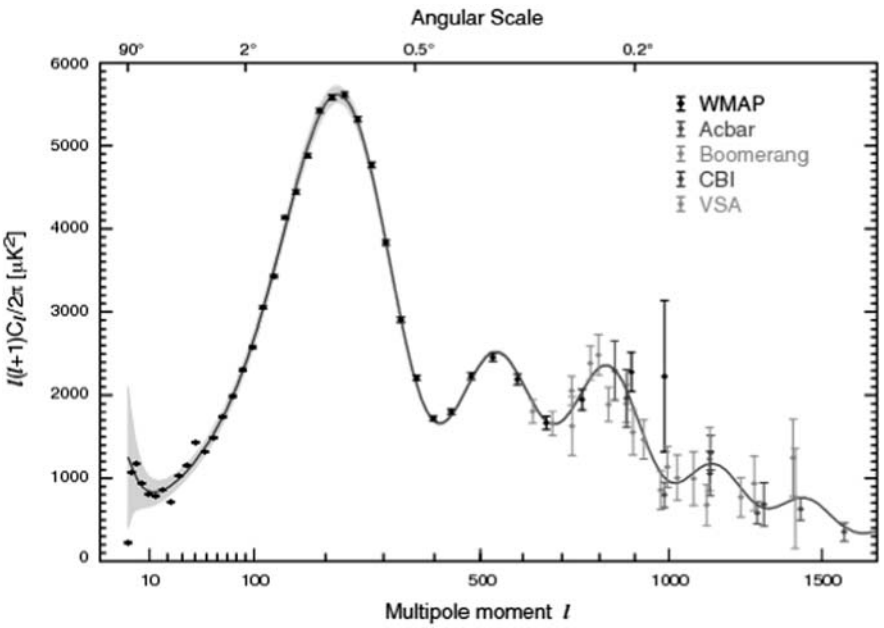

Fig. 3.9 Same as Fig. 3.8, but with additional data points from other observatories included. (Credit: NASA/WMAP Science Team)

spectrum following Eq. (3.30) and a Gaussian amplitude distribution. In Figs. 3.8 and 3.9 the brightness variations are expressed as temperature differences. This is possible since the CMB has a black-body spectrum, where the temperature provides an unambiguous measure of the local brightness and of the local energy density. Since the CMB brightness fluctuations are observed on the celestial sphere, their distribution is conveniently expressed in terms of spherical harmonic functions Y_{lm}. In this case the scale on the sky can be expressed by the index l of these functions. The corresponding angular scale on the celestial sphere is $\Theta = \pi/l$ (or, expressed in degrees, $\Theta = 180^0/l$). If a_{lm} are the amplitudes of the corresponding spherical components, the power spectrum of the observed fluctuations can be written as

$$C_l = \left\langle |a_{lm}|^2 \right\rangle \tag{3.34}$$

which is related to the Cartesian power spectrum $P(k) = Ak^n$ (Eq. 3.30) by

$$C_l = A2^n\pi^2 \frac{\Gamma(3-n)\Gamma\left(l + \frac{n-1}{2}\right)}{\Gamma^2\left(\frac{4-n}{2}\right)\Gamma\left(l + \frac{5-n}{2}\right)} \tag{3.35}$$

where Γ is the statistical Gamma function. For the special case of the Harrison-Zeldovich spectrum (n=1) C_l becomes

$$C_l = \frac{8\pi A}{l(l+1)} \tag{3.36}$$

Thus, for an unmodified Harrison–Zeldovich spectrum the function $l(l+1)C_l$ (plotted in Fig. 3.8) would be a constant.

As pointed out in Sect. 3.2.2, the observed CMB brightness distribution is a snapshot of the acoustic oscillations of the cosmic plasma at the time of decoupling of matter and radiation. Those density fluctuation which are observed in a phase of maximum contraction appear (as a result of the adiabatic heating) hot and bright in the microwave background. Correspondingly, a fluctuation which passes through a density minimum at the time of the CMB emission will appear cooler and fainter.[2]

Thus, those oscillations which at the cosmic time of the emission of the CMB had a maximal quadratic deviation from the mean density (i.e., either a peak or a valley of the time-variable density, causing, respectively, a bright or a dark spot in the CMB sky) produced maxima in the CMB power spectrum. The corresponding signal in the power spectrum is so strong since (as explained in Sect. 3.2.2) all oscillations of a given size or wavelength have the same phase when the CMB is emitted. On the other hand, for small fluctuations the signal is attenuated due to the fact that the CMB photons are emitted by a layer of finite thickness. This layer, from which we get the CMB photons, is called the "last scattering layer" or the "cosmic photosphere." From the baryonic density and ionization state at the corresponding epoch one can calculate that the comoving (FWHM) geometric depth of this layer is about 40 Mpc. Smaller fluctuations appear transparent and radiation from different phases becomes mixed. Therefore, the peaks from fluctuations with higher wave numbers become progressively smaller.

The most conspicuous feature of the CMB power spectrum (Fig. 3.8) is a dominant peak at $l \approx 220$. This (first) peak is due to the oscillation with the longest wavelength which by the time of the CMB emission could develop and which could reach a phase of maximum density. This condition means that the size of the corresponding fluctuation has to be about twice the size of the "sound horizon" $R_s = c_s t$, where c_s is the speed of sound, and t is the age of the universe at the time of the CMB emission. The comoving size of R_s is about 150 Mpc [47].

Since the first peak of the CMB power spectrum corresponds to a density maximum, it has been amplified by the gravitational dragging of the collapsing dark matter (see Sect. 3.2.2). The following peaks in the power spectrum correspond to fluctuations with sizes smaller by factors 1/2, 1/3, 1/4, etc. than the fluctuations causing the first peak. The positions of the peaks in the power spectrum follow approximately this sequence. (Their exact positions are influenced by the additional physical effects mentioned above).

---

[2] Additional effects influencing the CMB surface brightness are the Doppler effect resulting from the oscillations and the effects of the gravitational potential of the fluctuations on the photon energies. But these other effects contribute less the shape of the power spectrum.

While the first peak of the CMB power spectrum corresponds to a compressional phase (and to bright spots on the CMB sky), the second peak corresponds to a expansion phase (and dark spots on the CMB sky). Since the dark matter is collapsing, its dragging force attenuates the second peak, making it much weaker than the first peak. Analogously, all odd peaks (1st, 3rd, 5th, etc.), which correspond to density maxima, are amplified by the gravitational dragging, while all even peaks (corresponding to density minima) are attenuated. This odd-even effect is clearly visible in Figs. 3.8 and 3.9. In the absence of dark matter this odd-even effect would not be present. Thus, the relative amplitudes of the odd and the even peaks provide direct evidence for the existence of dark matter, and the quantitative evaluation of the ratio between the odd and the even peaks can be used to derive the baryonic matter fraction and the baryonic matter density at the time of the CMB emission. The baryonic matter density derived in this way from the CMB power spectrum is in excellent agreement with the baryonic density derived from the abundance of the light chemical elements (Sect. 3.1.4). Since the abundance of the light elements measures the baryon density at a redshift of $z \approx 10^9$, while the CMB power spectrum measures this quantity at $z \approx 1090$, the good agreement inspires some confidence in the current cosmological models.

Another important observed property of the CMB power spectrum is the angular scale corresponding to the acoustic peaks. Since the linear size of the corresponding fluctuations can be calculated from the physical processes described in the preceding sections, a comparison with the observed angular size can be used to derive the space curvature. The resulting curvature is consistent with k = 0 with a mean error of less than 0.04 [491].

The CMB power spectrum is sensitive to many other parameters determining the cosmological models. It has been measured rather accurately up to about $l = 1000$ (i.e., for comoving cosmic scales $> 50\,\mathrm{Mpc}$), making it the most important observation constraining the large-scale properties of the universe (see, e.g., [47, 491]). However, the sensitivity of the CMB power-spectrum fits for different cosmic parameters differs significantly, and the fits involve some level of degeneracy between individual parameters. Therefore, more stringent constraints and more reliable conclusions can be derived by combining the CMB observations with other tests. Particularly important are test observations pertaining to epochs of smaller redshifts, which are more sensitive to the dark energy, as the dark energy became dynamically important relatively late during the cosmic expansion.

3.2.3.2 Observations of the Baryonic Matter Distribution

As pointed out above, the cosmic density fluctuations with the largest amplitudes eventually resulted in the formation of galaxies and galaxy clusters. Therefore, in principle, studies of the large-scale distribution of galaxies also allow us to test the predictions of the theory of structure formation. This tests can be done at low redshifts and in the local universe. Moreover, studies of the galaxy distribution allow us to study the power spectrum of the fluctuations on relatively small scales where the present CMB data are less accurate.

The numerical simulations described above can be used to follow the evolution of the dark-matter fluctuations at least as long as a feedback of the baryonic matter evolution can be neglected. These computations make detailed predictions for the dark-matter distribution at the present epoch. On the other hand, the objects which we observe in the local universe are galaxies and clusters of galaxies. Because of complex physical processes involved in the formation of galaxies and clusters, it is not assured that the distribution of galaxies fully traces the distribution of the matter density. In particular, it is expected that the galaxy formation depends on the matter density in an nonlinear way. This effect is usually characterized by a "biasing factor" b, defined as the ratio between density contrast in the galaxy distribution and the density contrast in the dark-matter distribution or, equivalently,

$$b = \frac{\sigma_8(galaxies)}{\sigma_8(matter)} \tag{3.37}$$

If the galaxy formation follows exactly the dark-matter distribution, the biasing factor becomes $b = 1$. A biasing factor $b \neq 1$ is to be expected, when only mass concentrations exceeding a certain density result in the formation of galaxies.

Figure 3.10 shows a comparison of a (bias-corrected) power spectrum based on the galaxy distribution observed by the Sloan Digital Sky Survey (SDSS) [409] and the spectrum predicted on the basis of the CMB observations by the WMAP satellite. As pointed out in Sect. 3.2.2, the size dependence of the duration of the initial rapid ($\sim a^2$) growth of the fluctuation amplitudes results in a turnover of the present-day power spectrum at $k \approx 10^{-2}$ Mpc^{-1}. For larger values of k the theory predicts a rapid decrease of the power. Figure 3.10 obviously is in good qualitative agreement with this theoretical prediction. Figure 3.10 also shows that the observed power spectrum and the prediction based on the WMAP results agree well for $k < 0.07h$ Mpc^{-1}. But for larger wave numbers the power observed in the galaxy-density distribution is higher than predicted. This seems to indicate that nonlinear effects (which are particularly important on small scales) are not (yet) correctly taken into account in the models.

The density fluctuations which produce the acoustic peaks in CMB power spectrum also influence the dark-matter distribution and the accumulation of matter which form the galaxies. Therefore, the density variations which produce the acoustic peaks in the CMB power spectrum also leave traces in the power spectrum of the galaxy distribution. However, since the acoustic oscillations primarily affect the baryonic matter distribution, which in turn influences the dark matter, the amplitude of this effect (usually referred to as the "baryon oscillations") is small. It is detectable only in large data sets. So far the "baryon oscillations" have been detected in the 2dF Galaxy Redshift Survey [105] and (as illustrated by the inset of Fig. 3.10) in the Sloan Digital Sky Survey. As shown by Fig. 3.10 the observed baryonic observations are in remarkably good agreement with the prediction based on the CMB power spectrum.

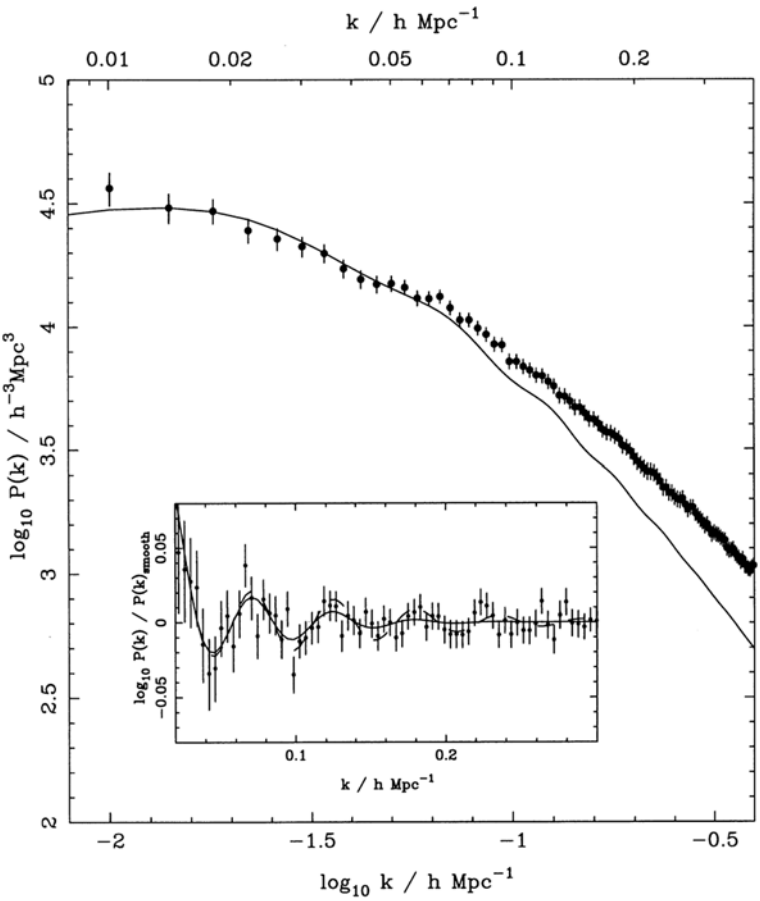

Fig. 3.10 Redshift-space power spectrum of the space distribution of $> 200\,000$ galaxies of the Sloan Digital Sky Survey (see Sect. 5.2.1). The *solid line* represents the galaxy power spectrum theoretically predicted on the basis of the (WMAP) CMB observations. The inset repeats the power spectrum near $k/h \approx 0.1$ normalized to the smoothed spectrum and on a different scale. Qualitatively, the shape of the power spectrum of the variations in the galaxy density is in good agreement with the theoretical expectation. Moreover, the baryonic oscillations (inset) are reproduced remarkably well. However, on small scales (i.e., at large values of k) the observed galaxy distribution shows more power than predicted. Since nonlinear effects are particularly important on small scales, this may indicate that these effects have not (yet) been adequately taken into account. (From Percival et al. [409], where details of the galaxy sample and the presentation are explained)

Instead of using the power spectrum, the baryon oscillations can also be detected by means of the spatial two-point correlation function $\xi(r)$ of the space distribution of galaxies. $\xi(r)$ is defined implicitly by the relation

$$N(r)\mathrm{d}V = N_0(1 + \xi(r))\mathrm{d}V \tag{3.38}$$

where $N(r)\mathrm{d}V$ is the number of galaxies in the volume element $\mathrm{d}V$ at a radial distance r from an arbitrarily selected galaxy, and where N_0 is the large-scale mean

number density. For a uniform distribution we have $\xi(r) = 0$. $\xi(r)$ obviously is a measure of the excess number of galaxies at a distance r relative to a uniform distribution. Since the power spectrum and the 2-point correlation function are connected by the relation

$$P(k) = V^{-1} \int \xi(r) \frac{\sin kr}{kr} 4\pi r^2 \mathrm{d}r \tag{3.39}$$

(where V is the volume over which the integration is carried out) $P(k)$ and $\xi(r)$ are essentially Fourier pairs. Therefore, the oscillations in the power spectrum result in a single peak in the two-point correlation function. This is illustrated by Fig. 3.11. The peak in the 2-point correlation function corresponds to the scale of the first maximum of the CMB power spectrum. The CMB observes this scale at $z = 1089$, while the galaxies are observed at much lower redshifts, or in the local universe. Figure 3.11 shows the correlation function (normalized for better visibility) for a sub-sample of the SDSS galaxies, again together with predictions based on the WMAP CMB data. The good agreement of the scales observed at very different redshifts confirms that our space is very close to being flat, and that the models are in agreement over a large range of redshifts.

As in the case of the CMB power spectrum, many different cosmological parameters can be constrained by the observed baryon oscillations, if the data are of sufficient accuracy (e.g., [155]). Relative to the observations of the CMB power

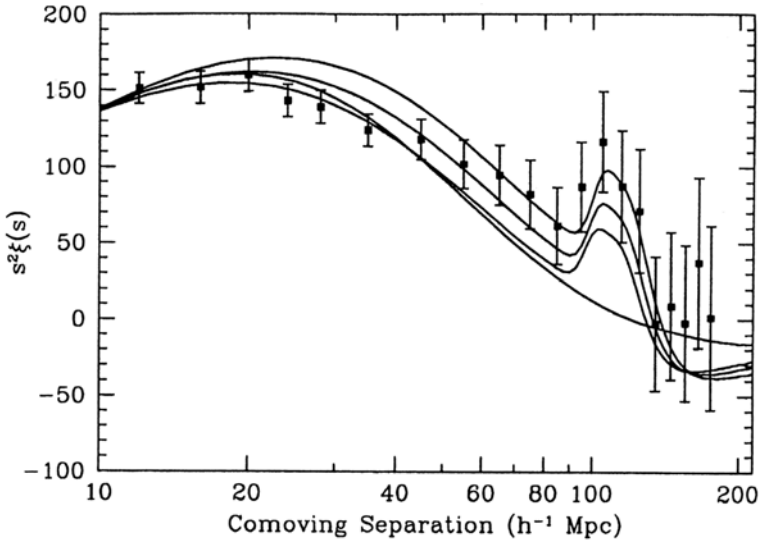

Fig. 3.11 Redshift-space correlation function of the SDSS Luminous Red Galaxy sample. The data points represent the observations. The three *upper lines* correspond to theoretical models consistent with the CMB power spectrum. The individual models assume slightly different matter densities. The (at the peak) *lower line* gives the predicted correlation function for the case that all cosmic matter is dark matter. (from Eisenstein et al. [155])

spectrum, the baryon oscillations have the fundamental advantage that such observations can be made for any redshift (or cosmic time) where adequate samples of galaxies can be collected.

In principle, the galaxy distribution allows us to test the fluctuation power spectrum at higher wave numbers than those observable at present in the CMB. However, at high wavenumbers the effects of the late, nonlinear evolution becomes important. These effects modify the power spectrum and eventually erase the original spectrum. An example of such nonlinear effects is the streaming of matter in the gravitational fields of galaxy clusters and galaxies. Such streaming of matter obviously re-arranges the cosmic density distribution. Therefore, the present galaxy distribution contains only limited information on the initial fluctuation spectrum on small scales. A better method to study the baryonic fluctuations on small scales is the analysis of the distribution of the intergalactic HI at intermediate redshifts. The properties of the HI distribution can be derived by analyzing the IGM Lyman absorption lines (the "Lyman forest") in the spectra of high redshift galaxies (e.g., [424]). The corresponding data available so far are consistent with the parameters obtained from the galaxy distribution and the CMB.

3.2.3.3 Weak and Strong Lensing

Another type of local or low-redshift observations providing information on the cosmic density fluctuations is an evaluation of the deflection and distortion of galaxy images by the gravitational field of the density fluctuations. This effect is known as "weak (gravitational) lensing" or "cosmic shear." Weak lensing has the same origin as the "strong" gravitational lensing effects observed near massive galaxies and galaxy clusters and described in Sect. 2.3. As illustrated in Figs. 2.29 and 2.30, strong lensing results in conspicuous arc-like distortions of the images of background sources. However, such strong lensing effects are observed only in the case of particularly dense galaxies or clusters. On the other hand, all mass concentrations cause low-level light deflections and distortions. These low-level effects are called "weak lensing."

Weak lensing by cosmic density fluctuations can be measured by determining the deviations of the distribution of galaxy shapes from a stochastic distribution. This can be done, e.g., by measuring the 2-point correlation function of the ellipticities of galaxy images. Since the weak lensing is caused directly by matter density fluctuations (which are dominated by the dark matter), observation of this effect provide the most direct and reliable method for deriving the fluctuation amplitude σ_8 of the matter. Recent results and their cosmological implications have been described, e.g., in [46, 186, 232, 540]. The value of the fluctuation amplitude derived in this way is in excellent agreement with the constraints provided by the CMB data. Since weak lensing and the CMB are observed at very different redshifts, this agreement again gives strong support to our present cosmological models.

Information on the density fluctuations in the local and in the low-redshift universe can also be derived from the observed frequency of the arc-like distortions caused by the "strong lensing" effects. The statistics of the strong lensing can be

compared to predictions of the occurrence of dense matter concentrations from the CDM simulations. However, in the particularly dense mass concentrations responsible for strong lensing, the distribution of the baryonic matter has to be taken into account. This complicates the interpretation of such data (see, e.g., [551]).

3.3 The Formation of the First Stars and Galaxies

3.3.1 Evolution of the Dark-Matter Halos

As pointed out in Sect. 3.2.2, part of the positive density fluctuations which survived up to the epoch of decoupling of matter and radiation, eventually stopped expanding, became gravitationally bound, and collapsed to form galaxies and galaxy clusters. A rough linear estimate ([404], see also Chap. 11 of [60]) indicates that only fluctuations which at redshift z_i had a density contrast of about

$$\delta\rho/\bar{\rho} > \frac{1-\Omega_M}{\Omega_M(1+z_i)} \tag{3.40}$$

can avoid expanding forever. From the analysis of the CMB data we know that at the epoch of decoupling only a fraction of the dark-matter fluctuations met this condition. Thus, it is not surprising that only about 8% of the baryonic matter ended up in the stars and the interstellar matter of present-day galaxies (cf. Sect. 2.7).

Other simple estimates show that the reversal of the expansion of the most dense fluctuations occurred after their density contrast $\delta\rho/\bar{\rho}$ had reached values of the order or above unity. Since these matter concentrations, which no longer participated in the cosmic expansion, consisted mainly of pressureless dark matter (and since the pressure of the baryonic matter was negligible at that stage) they immediately started to collapse. Part of the kinetic energy which the dark matter gained during the gravitational collapse was converted into random motions of the dark matter. These random motions prevented the matter from falling to the center of gravity of the fluctuation. Quantitatively, the end of the collapse of the dark-matter concentrations was determined by the condition that a virial equilibrium between the energy of the random motions E_{kin} and the gravitational energy E_{grav} was reached. Thus, the collapse was halted when $-E_{grav} = 2E_{kin} = 2E_{total}$. In a sphere of uniform density this happens after the radius has decreased by a factor 1/2. The real fluctuations were neither spherical, nor of constant density. However, more realistic numerical models confirm that the angular momentum stops the free collapse of the dark matter after a rather moderate density increase. As noted already, the resulting quasi-static configurations of virialized dark matter are called "dark-matter halos" or simply "dark halos" (as they were first observed as dark-matter halos around galaxies). According to the N-body simulations, up to the present epoch $> 50\%$ of the cosmic dark matter has been accumulated in such virialized halos.

Model computations show that dark-matter concentrations become bound halos when the density contrast relative to their environment has reached about a factor

of 100. Since a high contrast is reached first on small scales, the smallest matter concentrations which could collapse (i.e., which were more massive than the Jeans mass after the recombination) formed bound systems first. Therefore, the first dark-matter halos had mass values of the order $10^5 M_\odot$ to $10^6 M_\odot$ only. On larger scales the density contrast still continued to grow. As a result, the newly formed small halos started to cluster and to merge to form progressively larger bound systems. This hierarchical growth of the dark-matter halos is one of the most characteristic observable properties of the cold dark-matter cosmologies. As a result of the hierarchical merging and mass accumulation, the mass spectrum of the dark halos has been (and still is) changing with time.

As noted above, initially most the dark-matter halos were small, low-mass ($M \leq 10^6 M_\odot$) objects, and massive halos were very rare. Today the mass distribution is dominated by massive dark halos, while the number of low-mass halos is decreasing. Quantitatively, this evolution is shown by the N-body results plotted in Fig. 3.12.

A simple, but surprisingly successful analytic approximation of the hierarchical growth of halo masses has been derived by W. Press and P. Schechter ([421]). This formalism predicts the mass distribution of halos as a function of time. In its

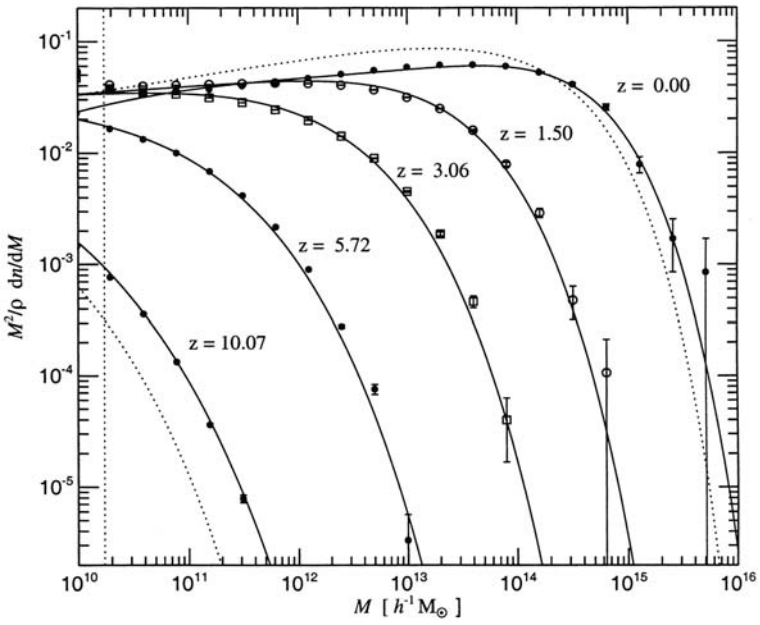

Fig. 3.12 Computed differential dark-matter halo number density distribution as a function of mass and redshift z. M is the halo mass and ρ is the mean cosmic density. The data points are derived from the Millennium Simulation, the *solid lines* are predictions from an analytic fitting function, and the *broken lines* are analytic predictions using the Press-Schechter model. The *error bars* represent the statistical uncertainties due to the finite number of halos in the simulation. (From Springel et al. [492], reprinted by permission from MacMillan Publishers Ltd from *Nature*, Vol. 435, © 2005)

modern form (derived by Efstathiou, [151]) the Press-Schechter mass function can be expressed as

$$N(M) \propto \frac{\gamma}{M^2} \left(\frac{M}{M^*}\right)^{\gamma/2} \exp - \left(\frac{M}{M^*}\right)^{\gamma}, \tag{3.41}$$

where N is the comoving number density of halos of mass M, $M^* = M_0^*(t/t_0)^{4/3\gamma}$ is a time-dependent characteristic halo mass, and $\gamma = 1 + n/3$, where n is the index of the original fluctuation power spectrum.

The Press-Schechter formalism also predicts that the radii of dark halos as a function of the redshift and halo mass M follow the relation

$$r_h \approx \left(\frac{G\,M}{100 H_0^2 E^2(z)}\right)^{1/3} \tag{3.42}$$

where G is the gravitational constant, H_0 is the Hubble constant, and $E(z)$ is the expansion function (see Equ. 3.17). (For a derivation of Eq. (3.42) see, e.g., [358, 359]). For a flat geometry and large redshifts z Eq. (3.42) approaches $r_h \propto (1+z)^{-1}$. The circular velocity of a halo of mass M and radius r is found to follow the relation

$$v_h = \left(\frac{G\,M}{r_h}\right)^{1/2}. \tag{3.43}$$

Results of a detailed numerical simulation describing the merging of small halos are presented in Fig. 3.13. In this simulation, which was carried out by T. Greif and colleagues [209], the dark-matter evolution was followed between the redshifts $z = 99$ and $z = 10$. Figure 3.13 shows the merger history of a halo which at $z = 10$ has reached a dynamic mass of $5 \times 10^7 M_\odot$. In the figure the formation of the individual small sub-halos and their merging to larger halos of increasingly larger mass is plotted as a function of cosmic time and redshift. As shown in Fig. 3.13, such diagrams resemble the hierarchical branching-out of a tree. Therefore, these diagrams are called "merger trees."

In the simulation of Fig. 3.13 the first resolved bound minihalo (with a mass of $2 \times 10^4 M_\odot$) was detected at a redshift of $z \approx 35$. By repeated small and large mergers eventually the mass of $5 \times 10^7 M_\odot$ at $z = 10$ was reached. As discussed below, at this epoch the gas in the halo had reached the conditions for efficient star formation.

As shown more clearly in Fig. 3.14, at the early epochs after the formation of the first halo, the formation of additional small halos results in a widening of the merger tree. At later epochs the merging of halos becomes the dominant evolutionary effect. At this stage the number of involved halos is decreasing again.

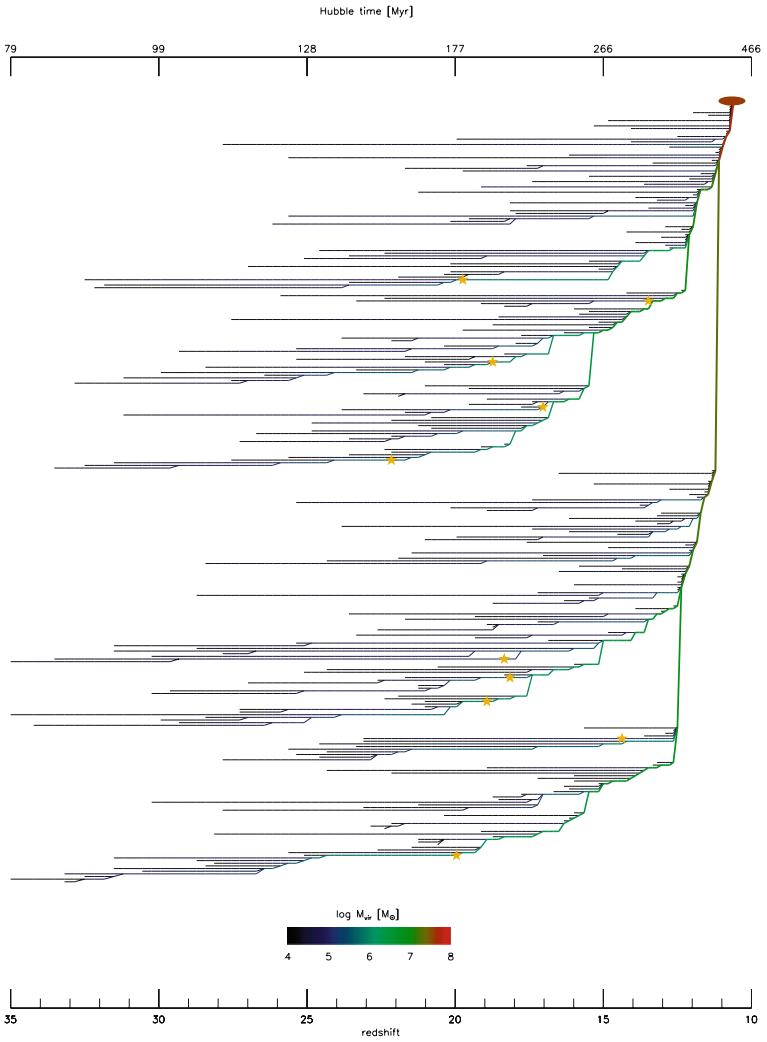

Fig. 3.13 Computed merger tree of the assembly of dark-matter halos resulting in a halo mass of $5 \times 10^7 M_\odot$ at $z = 10$. The abscissa gives the cosmic time (in Myr) and the redshift. The color code indicates the minihalo mass at the corresponding time. Small yellow stars indicate that the conditions for first star formation have been reached in the corresponding halo. (From Greif et al. [209])

3.3.2 Evolution of the Baryonic matter

Since there is much more dark matter than baryonic matter, the gravitational potential of the halos is dominated by the dark matter. The baryonic matter initially simply followed the dark matter to which it was bound by the gravitational field. As a result, initially the ratio between baryonic and dark matter in halos was about

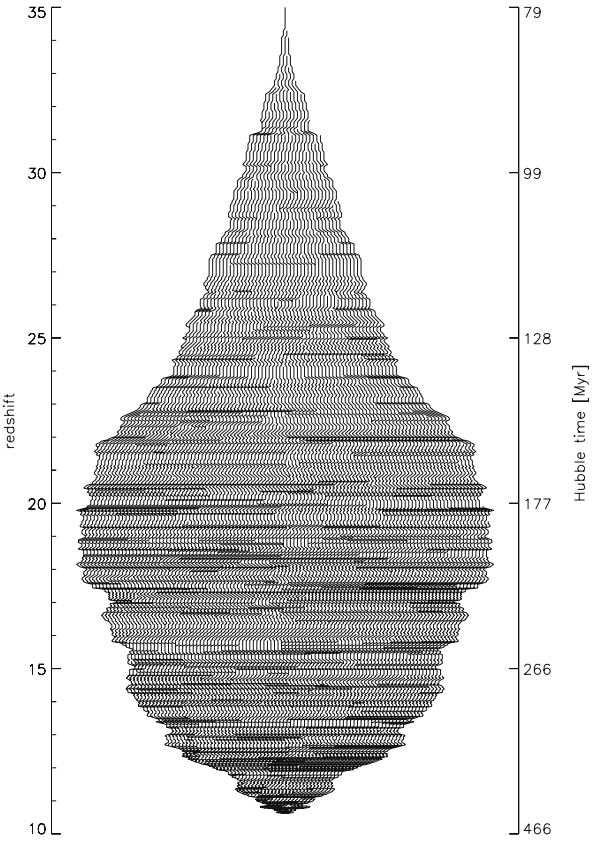

Fig. 3.14 Different presentation of the merger tree depicted in Fig. 3.13. As shown by the figure, at redshifts $35 > z > 20$ the rapid increase of the number of collapsing halos results in a widening of the merger tree. At later epochs the mergers start to dominate the evolution. As a result the merger tree becomes narrower again, until all the halos have merged into the single object which (in the simulation) is observed at $z = 10$. (From Greif et al. [209])

constant and close to the cosmic ratio ($\approx$ 0.18). However, since gas particles interact with each other, the mergers of the dark halos affect the dark matter and the gas differently. In particular, by producing pressure waves and shock fronts, the mergers result in the dissipation of mechanical energy and in a heating of the gas.

The gas can loose this thermal energy by means of inelastic particle collisions. Such collisions between the atoms, molecules, ions, and electrons result in the emission of radiation, which can escape and which results in cooling the baryonic matter. Since the angular momentum is conserved, while energy is lost, the cooling leads to rotationally flattened configurations and eventually to disks. A simple analytical theory for the disk formation, based on the analytic description of the dark-matter evolution outlined above, has been developed by Mo and colleagues ([358]). As shown

by these authors, if the surface density Σ of the disks can be described by an exponential

$$\Sigma(r) = \Sigma_0 \exp(-r/R_d), \tag{3.44}$$

and if the mass of the disk is a constant fraction m_d of the dark halo mass, and if, furthermore, the specific angular momentum of the disk is the same as that of the dark halo, the characteristic radius R_d becomes

$$R_d \propto r_h \propto m_d v_h E^{-1}(z). \tag{3.45}$$

Observations of local and low-redshift disk galaxies are in rather good agreement with this simple model.

In local galaxies (including our Milky Way galaxy) the cooling of the interstellar gas results from the collisional excitation of atoms and ions, followed by a radiative de-excitation. The radiation normally escapes from the cloud. The primordial gas clouds of the first low-mass dark halos consisted essentially of hydrogen and helium atoms, and the gas temperatures were typically $< 10^4$ K. At such low temperatures the electronic energy levels of hydrogen and helium atoms cannot be thermally excited. Thus the atomic collisions were essentially elastic. However, under the physical conditions of these early gas clouds occasionally H_2 and HD molecules were formed. Collisions of hydrogen atoms with these molecules can result in the excitation of rotational energy levels of H_2, followed by the emission of IR line radiation. Cooling by this process becomes important, when the fraction of hydrogen atoms bound in H_2 molecules reaches about 10^{-4}. Cooling by HD molecules is, in principle, more efficient, but these molecules are less abundant. The relative importance of the different cooling processes and the epochs where they become important has been investigated by various different groups, without reaching full agreement (see, e.g., [74, 254, 571], and Sect. 11.3).

Since the probability of particle collisions depends on the square of the particle density, the cooling of the gas by collisionally excited radiation increased rapidly with the increasing gas density. Where the cooling timescale became shorter than the free fall timescale of the gas, the cooling led to the the rapid formation of dense clouds and filaments of warm and cool gas (e.g. [570]).

Computer simulations of the cooling of the gas in the early dark-matter halos have been carried out by various different groups with different assumptions and different techniques (e.g., [1, 2, 72, 73, 209, 571]; see also the recent reviews in [74] and [254]). Although the results differ in details, the simulations show that the H_2 and HD cooling can lead to gas clouds with densities of the order 10^{-2} atoms m^{-3} and temperatures of about 200 K or lower. When these conditions are reached, star formation is expected to begin, and the baryonic mass concentrations inside the dark halos gradually become galaxies, i.e., bound configurations consisting of stars and interstellar matter. The current models suggest that the very first stars were formed at an epoch somewhere in the redshift range $20 < z < 30$. In Fig. 3.13 the points in time (and redshift) where the conditions for star formation are first met

are indicated by small yellow stars. As indicated by this figure, in the corresponding computer simulation the first stars formed in minihalos with masses between $10^5 M_\odot$ and $10^6 M_\odot$ at redshifts $z \leq 22$. However star formation took place in only a small fraction of these low-mass halos.

Although the first stars were formed by the same basic physical process (the gravitational contraction and collapse of a gas cloud), the initial conditions and the properties of the matter in the early dark halos were very different from the star formation regions in the local universe. The two most important differences were (i) the different initial chemical composition, and (ii) the different initial density and velocity distribution in the collapsing gas clouds.

As pointed out above, the clouds, in which the first stars formed, consisted of hydrogen and helium only, and their minimum temperature was determined by the temperature corresponding to the separation of the lowest H_2 and HD energy levels. In the local universe the interstellar gas can cool to much lower temperatures by exciting fine structure levels of atomic oxygen (O I), and singly ionized carbon (C II). Moreover, at high densities, the continuum emission of dust grains contributes to the cooling of the present-day ISM. In the primordial gas clouds these cooling agents were missing. Therefore, these clouds maintained a higher temperature, and the initial Jeans mass (Eq. 2.12) was correspondingly higher. The less efficient cooling of the primordial gas is also expected to have suppressed the fragmentation of the first collapsing gas clouds. Moreover, due to the different initial mass distribution, massive protostars could accrete more matter before the ignition of core hydrogen burning and the resulting stellar radiation pressure stopped the mass accretion. However, even in these first stars the accretion was eventually stopped by the feedback from the evolving hydrostatic core [342, 386, 509].

In the astronomical literature the first stars which formed from gas with an initial pure H-He chemical composition are denoted as "Population III" stars, to distinguish them from the Population I and Population II stars defined in Sect. 2.2.4.[3]

As a result of the higher initial gas temperature and the partial suppression of fragmentation, young Population III stars are expected to have higher masses than the young Population II stars. This is confirmed by model computations of such stars. However, the published models do not (yet) agree on the expected mass range of these first stars. Some calculations predict masses in the approximate range $100 < M/M_\odot < 500$ ([74], [387]), while other simulations result in more "normal" mass values of about 40 $M_\odot$. But all calculations agree that the first stars were massive, bright, hot, and short-lived, and that no low-mass Population III stars were formed. At the beginning probably only a few (and perhaps only single) massive stars were formed in a given low-mass dark-matter halo of about $10^6 M_\odot$.

Because of their high mass, the massive Population III stars predicted by the theoretical models only lived for about 10^6 years, before exploding as supernovae.

[3] Note that in contrast to the stellar populations I and II (which differ mainly with regard to their stellar mass and age) the Population III is defined exclusively by the absence of elements heavier than He in the initial chemical composition.

According to the predictions, part of these supernovae probably were pair-instability SNe (cf. Sect. 2.1.5), which resulted in a complete destruction of the stars and a dispersion of the stellar matter in space. Because of the (relative to present-day galaxies) rather shallow gravitational wells of the high-redshift, low-mass dark halos, the debris of these early SN explosions was spread over large volumes, resulting in an efficient and rapid chemical enrichment of the interstellar and intergalactic space.

With the beginning of the chemical enrichment of the interstellar matter the initial conditions for the formation of Population III stars changed dramatically, since already a small amount of heavy elements can lead to a more efficient cooling of the gas. Thus, the epoch of the massive Population III stars was probably a very short period, which soon was followed by the formation of normal (Population II) stars. According to the model computations by Bromm and colleagues [75], a carbon and oxygen abundance as low as 10^{-3} of the solar values lowers the typical mass of newly formed stars to the values observed in the local universe. However, in other simulations (based on somewhat different initial conditions) the effects of the changing chemical composition on the star formation was found to be less important [251]. The theoretical predictions are complicated by the fact that in addition to the formation of heavy elements their distribution in space has to be modeled correctly.

Apart from enriching the IGM with heavy elements, the first Population III stars also modified their environment by ionizing and heating their surroundings. This changed the conditions for the formation of the next generation of stars. On the one hand, the UV radiation from the hot Population III stars efficiently destroyed the hydrogen molecules which could cool the gas. On the other hand, after the Population III stars had ended their short life, the high free electron densities in the relic H II regions around them could assist the formation of new and H_2 and HD molecules. Therefore, the gas initially ionized by a Population III star could attain a higher molecular fraction and cool to lower temperatures. According to computer simulations, this resulted in a second generation of Population III stars with somewhat lower masses. In the literature these hypothetical second-generation stars are referred to as Population III.2 (while the very first population is denoted as Population III.1, see also Sect. 11.3).

The above scenario of initial generations of (only) massive stars (formed from H and He matter), followed by relatively normal star formation from slightly chemically enriched matter, appears to be consistent with the observational results collected in our Milky Way galaxy. Because of their short lifetime, the massive initial Population III stars have all disappeared. Low-mass ($M < M_{\odot}$) pure H–He stars, if they ever existed, could have survived. But among the very old, low-mass stars in our Milky Way galaxy, so far, no objects with a pure H–He initial composition has been found. This is consistent with the prediction that all Population III stars were massive. On the other hand, old low-mass stars with carbon and oxygen abundances approximately down to $\approx 10^{-3}$ of the solar abundance have been observed (e.g., [36]). This may confirm that a normal stellar mass function, including low-mass stars, was established as soon as the chemical enrichment of the interstellar matter had begun.

3.4 The Cosmic Future

As mentioned before, cosmological models can also be used to extrapolate the cosmic evolution into the future. Such extrapolations have been carried out to future epochs corresponding up to about 10^{100} times the present cosmic age (e.g., [134], [135], see also [344]). Although highly speculative, such studies provide interesting consistency checks for physical theories and concepts. Naturally, such extrapolations assume that the presently assumed laws of physics are correct and remain valid throughout the time range of the extrapolation.

Modern astronomical observations [436], [22], [491] indicate that since about 6 billion years our universe is expanding with an ever increasing rate. Therefore, the distances between the dark halos, between the field galaxies, and between the galaxy clusters will become larger at an increasing rate. At the same time the mean matter density drops $\sim a^{-3}$, and the temperature of the radiation field decreases $\sim a^{-1}$. Due to ongoing mergers of dark-matter halos and of galaxies the comoving number density of bound systems will become smaller, while the average mass of these systems is increasing. However, due to the increasing distances the merging processes will eventually come to an end and galaxies will stop growing, after having accumulated all matter available to them.

Today most of the cosmic matter density is due to the dark matter. The good agreement of the current cosmological models with the observations seems to indicate that the dark-matter particles can be assumed to be stable at least on a timescale of the present age of the universe. However, some theoretical concepts of the dark-matter particles suggest a long-term decay of the dark matter into fermion-antifermion pairs and/or radiation. This obviously would affect the dynamics of the bound systems in the universe. However, in the absence of detailed information on the nature and properties of the dark-matter particles, it is impossible to make meaningful predictions on the nature and timescale of such effects. Therefore, in the following the possible decay of the dark matter will be ignored and only the future of the baryonic matter will be discussed.

At present and in the near future the interstellar matter in galaxies is being progressively enriched with heavy elements through cycles of star formation and stellar mass loss. During this process, an increasing fraction of the matter is being locked up in black holes, neutron stars, and white dwarfs. The stellar-mass black holes will increase in numbers. The black holes in the centers of galaxies will grow in mass. However, the infall of matter into the supermassive black holes will decrease with time and will eventually fade out when all available material has been consumed. Thus, quasars and bright AGN will disappear. Together with the diminishing cosmic star formation rate this will result in a rapid decrease of the UV and FIR radiation fields and in a reduced heating of the remaining interstellar gas. As a result, dust formation in dense interstellar matter may become easier and the fraction of condensed cosmic mater will possibly increase with time, in spite of lower average gas densities.

The energy generation in stars requires nuclear fuel. As pointed out earlier, the most important nuclear fuel is hydrogen. Although at present the universe contains

plenty of hydrogen, its amount is decreasing and all hydrogen (and other nuclei providing a significant fusion energy potential) will eventually be exhausted. According to the estimates in the literature the last stars will stop producing nuclear energy when the universe has reached an age of about 10^{14} years (or about 10^4 times its present age). From then on only degenerate stars (white dwarfs, brown dwarfs, and neutron stars) will exist, which will cool to very low temperatures, and a new and final "dark age" will begin. As a result of this ultimate global energy crisis, life on planets in its present form will no longer be possible. However, planetary systems are expected to have disappeared by that time anyway, since on such long timescales close encounters of stars and planetary systems cannot be neglected. Such encounters tend to result in the ejection of planets (while captures are less likely). Simple estimates indicate that practically all planetary systems will be gone at a cosmic age of about 10^{17} yr.

An extrapolation beyond about 10^{33} years becomes uncertain because of the unknown long-term stability of the baryonic matter. Most of the mass in the cooling late universe, which has not been collected by black holes, will be present in the form of protons and neutrons. According to present experience protons (and neutrons locked in stable atomic nuclei) are stable. However, some versions of the GUT theories of particle physics suggest that protons actually decay. From experiments it is known that the mean life times of protons are at least of the order 10^{33} years. However, on longer timescales a decay of the protons cannot be excluded. If the protons decay, atomic nuclei and consequently the neutrons will also decay. Apart from photons and neutrinos the products of the decay will be electrons and positrons. Because of the (by then) extremely low cosmic density the positrons will survive. If protons decay, this will have a significant effect on the late cooling of the massive white dwarfs and the neutron stars. (Although the heating rate by proton decay would be extremely small, the very small surface of the more massive compact stars will result in an equilibrium temperature significantly above the surrounding radiation field.)

Finally, at a cosmic age of about 10^{66} years the stellar and at an age of about 10^{100} years the supermassive black holes will start decaying by means of Hawking radiation. The main product of this decay will be photons, which will disperse and disappear in the still expanding dark cosmos. Thus, if the assumptions made above are correct, very little will remain of the present, structured baryonic universe.

Chapter 4
Basic Techniques and Their Limitations

Practically, all information on the distant universe is based on the analysis of the electromagnetic radiation reaching us from high-redshift objects. In principle, relativistic particles, which are traveling at almost the same speed, neutrinos and gravitational waves can also convey information on distant galaxies. However, relativistic particles are normally charged. As a result, they interact with magnetic fields, with cosmic matter, and with the radiation field. These interactions tend to change their propagation direction, which makes it difficult to determine the origin of the observed relativistic particles. Neutrinos and gravitational waves propagate like light rays. But the low sensitivity of the corresponding detectors makes it impossible to observe high-redshift sources of neutrinos and gravitational waves.

The most distant observed galaxies and QSOs are intrinsically very powerful radiation sources. However, due to the large distance, only a small fraction of their radiation is intercepted by our telescopes. Moreover, the cosmic expansion dilutes and attenuates the small amount of energy which reaches us. Therefore, observing very distant galaxies requires techniques which allow us to detect and to evaluate extremely weak signals. This task is further complicated by the presence of foreground objects. At the signal level typical of the distant galaxies, anywhere on the sky there are many more (equally bright or brighter) foreground objects than high-redshift galaxies. Finally, for distant galaxies the surface brightness is comparable to, or lower than the sky background caused by the earth atmosphere (dominant in the case of ground-based observations), the interplanetary dust (which affects ground-based and space-based observations), and the galactic interstellar medium (affecting in particular the low-frequency radio observations). The following sections describe the instrumentation and the techniques which are used to meet this challenges.

Although (by detecting the distant "long" Gamma-ray bursts mentioned in Sect. 2.1.5) Gamma astronomy contributes to the studies of the high-redshift universe, most observations of high-redshift galaxies are made at optical (UV, visual, and IR), radio, and X-ray wavelengths. Therefore the following discussion will be restricted to these spectral regions.

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_4,

4.1 The Information Content of Light

Electromagnetic radiation is characterized by the direction from which the waves are reaching us, the radiation flux, the spectral energy distribution, and the polarization properties. In astronomy measuring the flux distribution as a function of the direction of incidence (or, equivalently, as a function of coordinates on the sky) is referred to as "imaging," while the derivation of absolute flux values is called "photometry." Because of the weak radiation fluxes, the accuracy of astronomical measurements is often limited by quantum effects. Energy received in the form of light always consists of a finite number n of photons with individual energies $h\nu$ (where h is the Planck constant and ν the radiation frequency). Since the arrivals of photons from astronomical sources normally are uncorrelated stochastic events, a measurement of n photons is subject to a statistical uncertainty of $\sqrt{n}$. Therefore, if n photons are observed, we have a statistical relative mean error of $\sqrt{n}/n = 1/\sqrt{n}$. Obviously, a photometric measurement cannot be more accurate than this limit determined by the photon statistics.

The surface brightness of high-redshift galaxies normally is lower than that of the night sky. As a result, the noise of photometric observations normally is determined by the photon statistics of the sky background, while the signal is proportional to the photon flux of the target. As discussed in detail in Sect. 4.2.4, the resulting signal-to-noise ratio of photometric measurements is in first approximation proportional to $\sqrt{\Delta t}$, where Δt is the integration time. As long as this approximation holds, with increasing exposure times progressively fainter objects can be detected. This also means that with longer integration times objects of constant intrinsic luminosity can be observed at larger distances. Therefore, long-exposure imaging observations, designed to reach very faint flux levels, are called "deep images."

Since $h\nu$ increases proportional to the frequency, the number of photons corresponding to a given amount of energy or flux becomes smaller with increasing frequency or decreasing wavelength. Thus, the photon statistics is a particularly serious limitation for X-ray and Gamma astronomy. With increasing wavelengths the photon statistics becomes less important. However, single photons can only be detected if their energies are larger than the thermal energy fluctuations kT of the detector or receiver. At radio wavelengths, where we typically have

$$h\nu < kT \tag{4.1}$$

(where k is the Boltzmann constant and T is the detector temperature) the error of the flux measurement normally is limited by these thermal fluctuations.

Although individual electromagnetic waves are polarized, the observed light from astronomical often shows no or little polarization since the superposition of many arbitrarily polarized waves results in a zero net polarization. Significant polarization is observed only if either the radiation source or the medium between the source and us is not (fully) isotropic. Therefore, the polarization state of the observed radiation contains information on the geometric properties and orientation effects of the sources and/or of the intergalactic and interstellar medium.

4.2 Imaging and Photometry

4.2.1 Telescopes and Interferometers

4.2.1.1 Single-Aperture Telescopes

The basic instrumentation used by astrophysicists consists of a telescope (which collects the light and sorts the photons according to the direction from where they are received) and a detector or receiver (which converts the incident radiation into an electronic signal). In many cases the instrumentation also includes a device which analyzes the radiation. This can be a filter, a spectrometer, a polarimeter, or any other device which allows us to determine the properties of the observed radiation.

The simplest telescope consists of an optical element or an optical system which images the sky on a detector. Many astronomical telescopes follow this principle. Modern large telescopes operating at wavelengths between the near-UV and the radio range normally use optics based on metallic reflection.[1] At shorter wavelengths, normally optical components based on grazing-incidence reflection are used. Since single-reflecting surfaces cannot produce good images of an extended field, in practice, reflecting telescopes consist of a large primary mirror, which forms the entrance aperture, and additional, smaller mirrors and/or additional lens optics. The first reflecting telescopes allowing to image large fields were developed early in the 20th century by the German optician Bernhard Schmidt, who combined a spherical mirror with an aspheric (4th order) large thin lense placed in the sphere's center. Schmidt-type telescopes played an important role in the photographic sky imaging and the first large area photographic sky surveys. Most present-day large astronomical telescopes use more complex and more efficient optical systems. A detailed description of the present-day technology of large optical reflecting telescopes can be found in [559] and [560]. Technical details of the reflectors used in radio astronomy are described in [25].

Because of the photon statistics, detecting and observing weak radiation fluxes from distant objects requires large light collecting areas. As a result, astronomical telescopes traditionally have apertures close to the limit which can be achieved with available technologies. But equally important for astronomical observations is the image quality produced by the telescopes. Among the criteria defining the image quality is the size and the shape of the "point-spread function" (PSF), which is

[1] For small optical telescopes refractive optics (lens systems) can also be used. However, with increasing lens size the light has to pass through a progressively thicker layer of the lens material. On the other hand, to get good images, the optical path lengths through the lens to the detector has to be constant within a small fraction of the wavelength. Since the optical path length in the lens is proportional to the index of refraction of the material, large lenses require very homogeneous glasses, which are difficult or impossible to produce. Moreover, the mounting of large lenses poses major mechanical problems. Therefore, lens telescopes are limited to apertures of about $\leq$1 m.

defined as the image produced by the telescope of a distant point source. As noted already, in order to obtain good images the phase differences between different beams reaching a given point in the focal plane must not exceed a small fraction of the wavelength. Therefore, mirror surfaces have to be accurate and smooth to a small fraction of the wavelength for which they are designed. Consequently, large mirrors producing good images are particularly difficult to achieve for short wavelength applications and the maximal feasible telescope aperture depends on the operating wavelength. The X-ray telescopes which are in use at present all have apertures <1 m. Optical single-aperture telescopes have diameters up to about 10 m, and the largest existing single-aperture radio telescope has a diameter of 305 m. Monolithic mirrors larger than about 8 m are difficult to transport and expensive. Therefore, telescope apertures above this size normally consist of accurately aligned mirror segments of smaller size. Lists and short descriptions of the most important currently used telescopes can be found in [42] (optical-IR) and [441] (radio) (see also Sect. 12.1.1).

In optical telescopes an image detector can be placed directly into the focal plane of the telescope, or the focal plane can be re-imaged on the detector. The re-imaging has the advantage that the effective focal lengths (and thus the image scale) can be optimized to fit the detector properties. Since at very large telescopes the image scale often is too large to match the detectors, re-imaging normally is used to reduce the effective focal length. Such re-imaging systems are called "focal reducers." The additional free parameters of the optics of the focal reducers (usually a collimator and a camera) can also be used to improve the image quality over the detector field. On the other hand, some light is lost in the additional optical components, and unsharp reflections from the focal reducer optics (which is located relatively close to the detector) can cause a non-uniform and variable background. Particularly troublesome is the so-called "sky concentration," which is light reflected from the detector itself and re-reflected back by the last optical elements.

Gravity forces, temperature variations, wind pressure, and other during astronomical observations unavoidable effects tend to produce deviations from the ideal mirror surfaces, which results in a degradation of the optical image quality. In order to obtain sharp images in spite of these effects, modern large telescopes use actively controlled mirrors where the mirror shape is constantly monitored (by means of wavefront sensors) and adjusted on a timescale of the order of minutes (e.g., Chap. 3 of [560]).

For measuring the wavefront aberrations in a telescope, a reference source producing a known wavefront pattern is used. Obvious reference sources are unresolved distant stars, which, in the absence of atmospheric aberrations, produce practically plane wavefronts. By comparing the observed wavefronts with the ideal wavefronts, the deviations caused by the deformed optics can be determined and the optics can be corrected accordingly.

From the standpoint of wave optics, a reflecting telescope produces an image of an object on the sky by means of constructive interference of the incoming electromagnetic waves. Therefore, the angular resolution of a telescope is limited by the diffraction pattern of a distant point source caused by the telescope entrance

aperture. For a circular aperture this diffraction pattern corresponds to an Airy function with the first minimum occurring at the radius

$$\alpha \approx \sin\alpha = 1.22(\lambda/D) \tag{4.2}$$

where λ is the wavelength used and D is the diameter of the aperture. Following a definition introduced by Lord Rayleigh in 1881, α is called the "resolution limit" or "resolving power" of a telescope with a circular aperture. Thus, at the wavelength of 1 μm the diffraction-limited resolving power of a 10-m telescope corresponds to about 0.025 arcsec. At 10 μm the resolution limit, according to the Rayleigh criterion, is 0.25 arcsec. For single-aperture radio telescopes the ratio λ/D typically is 10^3 times larger and the resolution limits normally are > 10 arcsec. The resolution of current X-ray telescopes is limited by the aberrations of the optics or by the detector resolution.

For two point sources with an angular distance corresponding to the Rayleigh resolution limit, the first minima of the Airy functions coincide with the maximum of the neighboring object. In the absence of other effects, the flux minimum between the two (equal) point sources corresponds to 71% of the maximum flux. With good S/N two point sources can be distinguished even at distances somewhat below the Rayleigh resolution limit, but for poor S/N data the actual resolution limit may be lower than expected from the Rayleigh resolving power.

4.2.1.2 AO Techniques

Actively controlled optical elements can also be used to compensate the wavefront aberrations caused by the refractive-index variations which are produced by the turbulent Earth atmosphere. The refraction effects caused by the atmospheric turbulence elements result in a widening of the effective PSF and (due to random interference effects) in short-term random variations of the flux of unresolved astronomical objects. In the astronomical literature the widening of the PSF is referred to as the "seeing," while the flux variations caused by the turbulent atmosphere are called "scintillation." The scintillation and seeing effects decrease with increasing wavelengths. Therefore, the atmospheric effects are negligible for radio observations. On the other hand, observations at radio wavelength are affected by a similar effect occurring in the interstellar plasma. (The visual light is not affected by the ISM.)

Since the atmospheric seeing and scintillation at optical wavelengths are caused by the motions of the turbulence elements in the Earth atmosphere, these effects vary on the timescale of the motion of these elements through the line of sight. This timescale is of the order 10 ms, which is much shorter than the time on which the large and massive primary mirrors of astronomical telescopes can react. Therefore, for compensating the atmospheric wavefront aberrations additional optics has to be added where an image of the entrance pupil is projected on a small active mirror which can be modulated on the required short timescale. Such active optical devices,

which can correct the atmospheric wavefront aberrations, are known as "adaptive optics" or "AO" systems (for details see, e.g, [440]).

The precision to which wavefronts can be measured is limited by the photon statistics of the observed light. Thus, the accuracy depends on the product of the photon flux received from the star and the integration time. For the correction of the slowly variable deformations of the telescope components caused by gravity and temperature effects, the photons can be collected over typical integration times of minutes. Therefore, relatively faint reference stars can be used for measuring and correcting the aberrations caused by the optics. Such reference stars can always be found near an astronomical target. The situation is less favorable for the correction of the atmospheric aberrations. Because of the short timescales of the atmospheric variations, in this case the integration times are limited to milliseconds. Thus, much brighter reference stars are required, which may not be present near a target. In this case adaptive optical systems have to make use of "artificial stars," which can be generated by the back-scattering of laser beams focused to the high atmosphere (Fig. 2.6). At present such "artificial stars" are in most cases produced by means of lasers emitting light at the wavelengths of the resonance doublet of atomic sodium (e.g., [519]). In the Earth's atmosphere atomic sodium is enriched in a relatively thin layer at a height of 90–100 km above the Earth's surface. (Above this layer sodium is ionized, in the lower atmosphere sodium normally is bound in molecules.) Thus, the distance of artificial stars based on the sodium resonance scattering is relatively well defined.

Artificial stars can also be produced by Rayleigh scattering of laser beams. Since Rayleigh scattering occurs at any height in the atmosphere, in this case pulsed lasers have to be used, and the light backscattered from a certain height has to selected by making use of the height-dependent time delay of the optical echo.

Since light rays reaching a telescope from different points on the sky pass through different regions of the atmosphere, adaptive optical systems using a single reference star can correct the atmospheric aberrations only in a small area around the reference object. In the visual light the radius of this "isoplanatic" region is less than one arcmin. A correction over a larger field is possible, if an adequate number of reference stars, distributed over the field, is used. In this case the wavefront aberrations have to be measured separately for each reference star. With several reference stars it is also possible to separate the wavefront aberrations produced in different atmospheric layers and to correct the effects of the individual layers using several adaptive optical components. However, with increasing field size and the correspondingly increasing number of reference stars such "multi-conjugated" AO (MCAO) systems become increasingly complex (for details see, e.g., [437]). Therefore, the angular resolution of ground-based wide-field images normally is still limited by the seeing. The corresponding PSF varies with the time and the location. An example for the distribution of the mean FWHM of the seeing PSF at a typical modern observatory site is given in Fig. 4.1. The numerical values given in Fig. 4.1 correspond to observations at $\lambda = 500$ nm made near the zenith. For other wavelengths and for zenith distance angles $z > 0$, the seeing can be estimated using the approximate relations $\mathrm{FWHM} \propto \lambda^{-0.2}$ and $\mathrm{FWHM} \propto (\cos z)^{-0.6}$.

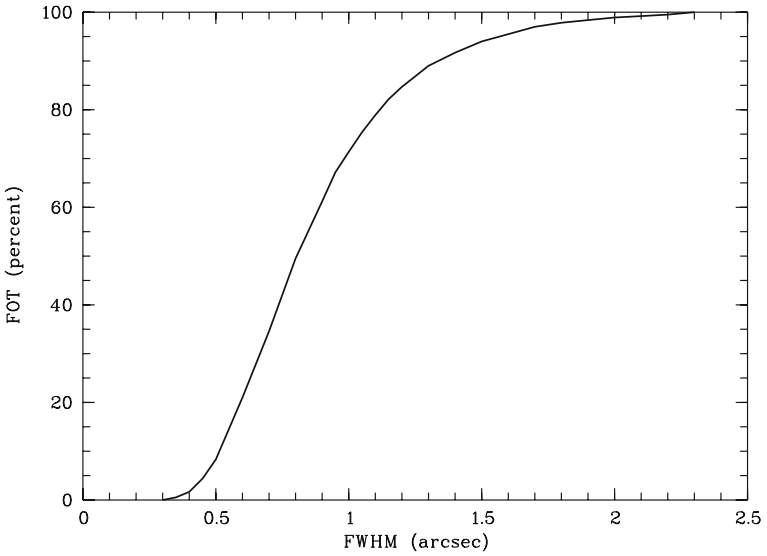

Fig. 4.1 Accumulated seeing distribution for 10-minute optical exposures at the ESO VLT Paranal observatory. Plotted is the fraction of time (FOT) where the seeing at $\lambda = 500$ nm was better than the given FWHM value during the 5-year period from 1999 to 2004. (Based on information presented on the Paranal Observatory Web page)

4.2.1.3 Space-Based Telescopes

Most of the electromagnetic radiation reaching us from the distant universe cannot penetrate the Earth's atmosphere. Only for the visual light, for part of the infrared, and for most of the radio range our atmosphere is more or less transparent. Telescopes operating not in one of these atmospheric windows obviously have to be based outside the atmosphere. Suitable platforms for space telescopes are artificial Earth satellites, artificial planets, or (in the future) the Moon.

While astronomers observing at wavelengths outside the atmospheric windows have no choice, observations at optical and radio wavelength can be carried from the ground or from outside the Earth's atmosphere. Observations from space always have the significant advantage that the absorption, scattering, and refraction effects of the Earth's atmosphere are absent. Therefore, space-based optical telescopes can produce diffraction-limited images without adaptive optics. Moreover (since the contributions from natural and man-made sources in the Earth's atmosphere are absent), in space the sky background is lower, which is particularly welcome for observations of very faint objects, such as distant galaxies. As a result of this advantage, many of the most distant known galaxy samples have been identified using the Hubble Space Telescope (HST), although the 2.5-m HST has only about 6% light-collecting power of the existing ground-based 10-m telescopes. On the other hand, the construction and operation of large space-based telescopes is much more expensive than in the case of ground-based facilities. As a result, observing time on

space facilities is precious and usually allocated in small portions. Moreover, with the currently available technologies much larger telescope apertures can be realized on the ground than is possible in space. Therefore, applications where high photon fluxes and/or large amounts of observing time are needed, are still the domain of ground-based telescopes, while space instruments have clear advantages for the diffraction-limited imaging of faint objects.

Among the fields where space-based telescopes have less advantages is spectroscopy in the optical bands. In this case observations with large ground-based telescopes normally give a better S/N since more photons can be collected. In the red spectral range, ground-based spectroscopy of faint objects normally is limited to the spectral regions between the atmospheric OH lines. Since the atmospheric OH lines are sharp, the amount of spectral information which is lost by superimposed OH features can be minimized by choosing an adequately high spectral resolution.

Fig. 4.2 The Spitzer infrared space observatory during the preparation for launch at NASA. The image shows the cooled 85-cm telescope on top of the instrument assembly. In space the telescope is shielded from the solar radiation by the solar panels visible behind the telescope tube. (Photo National Aeronautics and Space Administration)

At infrared wavelengths space-based observations practically always give better results, since the limited IR transparency and the thermal emission of the atmosphere reduces the S/N of ground-based observations for imaging as well as for spectroscopy. Moreover, because of the less energetic IR quanta, photon noise is less critical at long wavelengths. Therefore, at IR wavelengths cooled space telescopes with small apertures often are more sensitive than large ground-based telescopes.

Of particular importance for the recent progress in the field of high-redshift galaxies has been the NASA Spitzer Infrared Space Observatory (Fig. 4.2). With Spitzer it became possible to carry out photometry and spectroscopy of faint sources in the wavelength range 3 μm to 180 μm. In this way it became possible to study the evolved stellar populations and passively evolving galaxies at high redshift, as well as the emission of warm dust in high-redshift starburst galaxies.

4.2.1.4 Interferometers

As noted above, in the focal plane of a telescope, images are formed by the interference of all waves reaching the telescope aperture. Therefore, in principle, images can also be derived by collecting the light at different points or areas of a (real or virtual) aperture and by combining the corresponding light beams coherently. This technique is called "aperture synthesis" or "astronomical interferometry." It is particularly important for imaging in the radio range, where high-resolution observations with single apertures are not possible since the angular size of the diffraction pattern is to large (cf. Eq. 4.2). But interferometry is also an important tool to map the spatial fine structure of astronomical objects and to measure accurate positions at IR and visual wavelengths.

In practice, interferometers make use of the fact that the information contained in an image can either be presented by the image itself or by its Fourier transform. Mathematically, the image is given by the observed radiation intensity as a function of coordinates on the sky, i.e. by the function $I(\alpha, \delta) = I(\boldsymbol{a})$, where α and δ are coordinates in an image plane or on the sky, and $\boldsymbol{a}$ is a vector representing these two coordinates. The Fourier transform $F(\boldsymbol{u})$ of the image is given by

$$F(\boldsymbol{u}) = \int I(\boldsymbol{a}) \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{a} \cdot \boldsymbol{u}} \mathrm{d}\boldsymbol{a} \tag{4.3}$$

where **u** is a vector representing coordinates u and v in the spatial frequency space (e.g., [313]). If one of the two functions I and F is known, the other one can be calculated by means of a 2-d Fourier transformation (or inverse Fourier transformation). Obviously, the image function I is always a real function. Its Fourier transform F normally is a complex function. The imaginary part of F is zero only if I is completely symmetric (with respect to the image center) in both coordinates. Astronomical interferometers basically measure the Fourier transform F of an intensity distribution on the sky (or, at least, parts or properties of the function F). They make use of the fact that F is closely related to the mutual coherence function (as defined

in theoretical optics, e.g., [59]) of the light entering a (real or virtual) telescope aperture.

The most simple interferometer consists of two telescopes looking toward the same direction on the sky, with the two beams having a projected distance d (measured perpendicular to the observing direction, as indicated in Fig. 4.3). If combined coherently, the waves reaching us from the part of the sky which corresponds to the field of view (FOV) of the individual telescopes, are amplified or attenuated, depending on the phase shifts between the two beams. From simple geometric considerations it is clear that these phase shifts depend on the direction to the radiation source, on the wavelength, and on d. If the observed object is a point source, for a given wavelength a variation of d or of the observing direction results in a periodic series of maxima and minima, which are known as "interference fringes." If the observed object has (over the FOV of the individual telescopes) a uniform surface brightness, the signal resulting from the beam combination is independent of d and the observing direction. In this case no fringes are observed. Obviously, the presence and amplitude of a fringe pattern depends on the intensity distribution $I(\alpha, \delta)$ of the object on the sky. The amplitude of the fringes as a function of d is called the "fringe visibility" or simply "visibility" $V(d)$. This function obviously contains information on the size and the structure of the observed object.

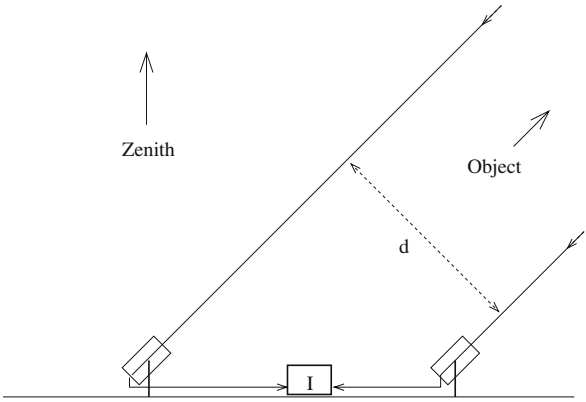

Fig. 4.3 Schematic outline of a 2-telescope interferometer

From the geometry of the 2-element interferometer it is obvious that any displacement of a resolved image feature relative to the image center results in a phase shift of the fringes relative to the fringe pattern that is (or would be) produced by a point source at the image center. Therefore, to fully characterize the "fringe visibility," one needs two real functions (the fringe amplitude and the phase shift), both being functions of d. These two real functions can be combined to form one complex function, which is called the "complex visibility" $V_c(d)$. Except for the special case of a rotationally symmetric images, the complex visibility function depends on the length of the projected distance d as well as on the orientation of the line separating the two beams. Thus, the visibility actually is a function of a vector $\boldsymbol{d}$ which represents the linear distance d and its direction. As shown in

the textbooks on observational astronomy (e.g., in Chap. 8 of [441]), the complex visibility function $V_c(\boldsymbol{d})$ corresponds (apart from a normalization factor and a factor due to the diffraction effects of the individual telescopes) to the mutual coherence function mentioned above. Therefore, by means of a Fourier transformation $V_c(\boldsymbol{d})$ can be converted into the image function $I(\boldsymbol{a})$, if the properties of the interferometer and the properties of the individual telescopes are known.

In principle, the integrals of the Fourier transformations have to be calculated over the complete coordinate space of the independent variables. Consequently, for a complete image reconstruction the complex visibility function has to be known for all values of $\boldsymbol{d}$, i.e. for $-\infty < u < \infty$ and $-\infty < v < \infty$. At any given time a 2-element interferometer obviously measures the visibility only for one particular value of the vector $\boldsymbol{d}$. The visibility for additional values of $\boldsymbol{d}$ can be obtained by varying the distance or the relative orientation of the telescopes, or by adding more telescopes at different distances and with different orientations of $\boldsymbol{d}$ to the interferometer. At astronomical interferometers this is somewhat simplified by the fact that, due to the Earth's rotation, during an observation the zenith distance angle ζ of a target and the orientation of the telescope baseline relative to the sky normally are changing. Since $d \propto \cos\zeta$, even at a 2-element interferometer with fixed telescopes the projected distance d and its orientation automatically cover a certain range during an astronomical observation. A full coverage of the u–v plane can be achieved if at least part of the telescopes can be moved. In this case a full sampling of all baselines within are virtual aperture corresponding to the size of the interferometer is possible. But a full sampling is time consuming. Therefore, even at large, many-element interferometers normally only incomplete complex visibility functions are produced. In order to obtain images from such incomplete visibility data, sophisticated iterative procedures have been developed. These procedures use model fits to constrain the image properties and to derive images which at least correspond closely to the correct image functions. Details of the currently used procedures and references to the extensive literature on this subject can be found in [441].

At radio wavelength interferometers today routinely produce excellent images of comparable or better quality than direct sky images obtained at optical wavelengths. The most important current radio interferometer for weak astronomical sources is the "Very Large Array" (VLA) at Socorro, New Mexico, USA (Fig. 4.4). This interferometer consists of 27 movable radio telescopes of 25 m diameter each, which can be configured in different patterns with maximum distances up to 25 km. An even more powerful radio instrument will be the "Atacama Large Millimeter Array" (ALMA), which is under construction at Chajnantor in northern Chile. ALMA consists of 64 telescopes of 12 m diameter. While the VLA operates mainly at centimeter wavelengths, ALMA is designed for observations at wavelengths between about 0.3 mm and 1 cm.

Although important results (such as surface maps of stars) have already been obtained, interferometric imaging at IR and optical wavelengths is still a developing field. Because of the shorter wavelengths, accurate (fringe) phase measurements are more difficult in the optical and IR than at radio wavelengths. Therefore, observers using optical interferometers often have to content themselves with deriving the real

Fig. 4.4 The Very Large Array of the National Radio Astronomy Observatory at Socorro, New Mexico. (Image courtesy NRAO/AUI)

visibility $V(\boldsymbol{d})$ (i.e., the modulus of the complex visibility). While no true images can be derived from $V(\boldsymbol{d})$ alone, this function already provides direct information on the extent of a simple object along the direction of d. Thus, if $V(\boldsymbol{d})$ is measured for several directions of $\boldsymbol{d}$, the shape and the size of simple objects (such as rotating stars or disks) can be determined.

An example of a large optical interferometer designed for the high-resolution imaging of faint targets is the "Very Large Telescope" (VLT) of the European Southern Observatory (ESO) at Paranal, Chile (Fig. 4.5). The VLT consists of four stationary 8-m telescopes and four movable 1.8-m telescopes. The light from these telescopes can be combined coherently in a central interferometric laboratory. While the four large VLT unit telescope are also used as single instruments, the four 1.8-m telescopes are dedicated exclusively to interferometry.

4.2.2 Image Detectors

The most basic method of obtaining information on the distant universe is taking deep pictures of the sky. Many properties of the recorded astronomical objects can be derived by analyzing such images.

The first astronomical images were drawings based on visual observations carried out with the naked eye or a telescope. Starting with first photographs of the Moon (taken in 1841), for more than a century photographic plates were the main imaging detectors in astronomy, before being replaced by more sensitive and more convenient electronic imaging techniques.

Fig. 4.5 The Very Large Telescope of the European Southern Observatory. The VLT consists of four large (8.2 m aperture) stationary telescopes and four smaller (1.8-m aperture) movable "auxiliary" telescopes. The 8.2-m telescopes are housed in the large box-like enclosures. One of the auxiliary telescopes (in a smaller, dome-shaped enclosure) is visible to the right of the image center. In the foreground the tracks are visible on which the auxiliary telescopes can be moved to different positions. For interferometric observations light from the individual telescopes can be combined in the flat building between the auxiliary telescope and the rightmost large telescope. (Photo: European Southern Observatory)

Today images of the sky are routinely obtained in many different wavelength bands throughout the electromagnetic spectrum. At short wavelengths (i.e., at Gamma-ray, X-ray, optical, and infrared wavelengths) normally imaging photon detectors are used. At radio frequencies, where (as outlined in Sect. 4.1) the detection of individual photons is not practicable, coherent radiation detectors (or "receivers") are used. A third type of astronomical detectors are the bolometers, which measure radiation indirectly by means of the temperature increase of a suitable radiation absorber. In principle bolometers can be used throughout the electromagnetic spectrum. But their broad sensitivity range make bolometers particularly susceptible to background effects. In modern astronomy bolometers are mainly used as broad-band detectors of millimeter and sub-millimeter waves (e.g., [23, 418]).

Most current photon detectors for the X-ray, optical, and infrared spectral ranges make use of the internal photo-effect in semiconductors. At present, the standard detectors for imaging at X-ray and optical photon energies are silicon-based charge-coupled devices (CCDs). For the detection of infrared photons with energies below the (1.1 eV) band-gap energy of silicon, diode arrays of materials such as Ge, InSb, HgCdTe, or doped Si or Ge impurity-band semiconductors are used. These IR-sensitive diode arrays usually are bonded or "hybridized" to "readout integrated circuits" (ROICs), which use conventional complementary metal oxide semiconductor ("CMOS") technologies. Detectors based on this technology are referred to as "hybrid detectors" or "CMOS detectors." Since the ROICs normally include a separate amplifier for each pixel, CMOS detectors allow more flexible operation

modes than CCDs. With suitable diode materials (such as Si) CMOS detectors can also be used in the visual and UV spectral range. But, so far, hybrid detectors are more expensive than CCDs of comparable performance. A comprehensive account of the present state of astronomical CCD and CMOS detectors can be found in [38]. A comprehensive overview of current IR detectors is given in [435].

Well-engineered astronomical photon detectors have a quantum efficiency QE (defined as the fraction of the incident photons which is actually detected by a given device) above 60%. Good detectors reach peak QE values $> 95\%$.[2] Thus, in a good and well-matched astronomical photon detector, practically each incident photon is converted into a photo electron and recorded as a signal. The main intrinsic noise source of modern array detectors is the "readout noise," which is due to stochastic variations of the charges which (even in the absence of photoelectrons) are present at the input ports of the on-chip amplifiers. Depending on the type of device, the readout noise may range between about one electron and more than 100 electrons. Because of the small capacities of their on-chip amplifiers, at present, CCDs tend to have the lowest readout-noise values.

During an exposure the photoelectrons generated in CCDs and diode arrays are kept in position by local electrostatic fields. This confining field is weakened by the negative charges of the accumulating photoelectrons. Therefore, for each picture element (or "pixel") there exists a well-defined upper limit for the number of photo electrons which can be accumulated during an exposure. This upper limit, where a pixel becomes saturated, is called the "well capacity" of the device. For astronomical CCDs designed for the optical spectral range, the well capacity typically is of the order 10^5 electrons. The ratio between the well capacity and the readout noise defines the detector's dynamic range (i.e., the flux ratio between the brightest and the faintest source which can be observed during the same exposure). The exact value of the well capacity depends (among other parameters) on the pixel size. Larger pixels result in a higher well capacity. On the other hand, the physical size of CCDs and integrated diode arrays is limited by the size of the largest semiconductor wavers which can be manufactured of the corresponding material. Therefore, with small pixels a larger total number of picture elements can be accommodated on a detector chip and a higher image resolution can be achieved.

Obviously, the well capacity limits the integration time for a single exposure. In the case of faint distant galaxies, CCDs and IR detectors become saturated by the flux from the sky background long before enough photons from the galaxies can be collected. In the blue and visual spectral range the maximal exposure times for broad-band photometry typically are of the order of a few minutes. In the NIR spectral range the stronger sky background limits the exposure times to seconds. For imaging through narrow-band filters and for slit spectroscopy, longer exposure times are possible. But in any case, the long integration times required for detecting and measuring faint objects have to be split into many shorter unsaturated exposures.

---

[2] For comparison, photographic emulsions normally have quantum efficiencies $\leq 1\%$, while photocathodes have typical QE values of the order 10–30%.

These short exposures are then co-added or "stacked" to generate the final images. To simplify the "flatfielding" procedure (see Sect. 4.2.3) and to reduce the effects of bad detector pixels, the telescope is slightly moved between the individual exposures, so that on the individual frames the observed objects are recorded by different detector pixels. The technique involving the movement of the telescope between the individual exposures is called "dithering."

Currently the most common imaging detectors for the far infrared, sub-millimeter, and millimeter wavelengths are the bolometer arrays. At noted above, bolometers measure radiation fluxes indirectly by the temperature increase generated in an absorber. In current bolometer arrays the temperature increase is measured using either temperature-sensitive semiconductors (e.g., [280, 478]) or temperature-sensitive superconducting devices, such as transition edge sensors (TESs) (e.g., [23]). TES bolometers are technically more complex, but more sensitive. Moreover, using photo-deposition technologies, larger pixel numbers can be achieved with TES bolometers.

At longer radio wavelengths the low angular resolution makes imaging with single telescopes impracticable. Instead, interferometers with coherent radio receivers at the individual telescopes are used to obtain image information.

4.2.3 Image Reduction and Image Analysis

4.2.3.1 The Treatment of Raw Images

The raw output of imaging detectors normally consists of a digitally recorded voltage or charge for each image pixel. Within the dynamic range of the detector this voltage or charge is proportional (or otherwise uniquely related) to the integrated flux received by the pixel. Since the surface brightness of distance galaxies generally is below that of the sky background, and since all detectors produce systematic errors, the raw images normally are dominated by the background and by instrumental effects which often completely hide the astronomical signal (e.g., Fig. 4.6). In order to extract the signal, the raw images have to be treated using suitable reduction procedures.

For technical reasons, the raw signal often includes an additional (approximately constant) voltage, which is called "bias." As a first step, the bias, which may be position-dependent, has to be determined and subtracted using unexposed pixels or dark exposures.

Since the quantum efficiency of a detector varies between individual pixels, the next step of the reduction is a careful correction of these pixel-to-pixel variations. For this purpose, images of a flat intensity distribution are obtained and the science images are divided by these "flat-field images" (which in the astronomical literature are referred to as "flats"). If carried out properly, this procedure also removes large-scale sensitivity gradients due to optical vignetting and field-dependent transmission variations of the optics. Since the surface brightness of distant galaxies is well below the sky background, even small errors in the flatfielding procedure can

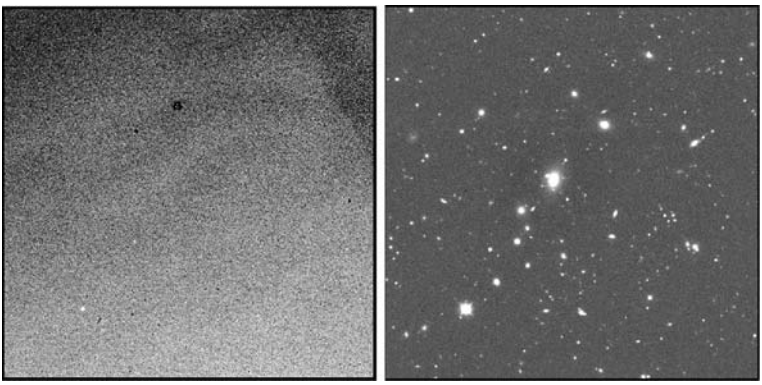

Fig. 4.6 Comparison of a raw and a reduced image obtained with an imaging solid-state detector. The *left panel* shows an individual raw image obtained with an HgCdTe detector in the H-band ($\lambda = 1.63\,\mu$m). The *right panel* shows the same field after image reduction and co-adding of 50 dithered individual images. (From Gabasch et al. [190])

introduce an amount noise which is large compared to the signal and which can render the observations completely unusable. Since the quantum efficiency depends on the wavelength, the light source used for flatfielding of sky images has to have the same spectral energy distribution as the night sky.

For flatfielding ground-based optical observations, normally images of the twilight sky are taken before and after the observing nights. Flats obtained in this way often give satisfactory results since the color of the twilight sky is similar to that of the night sky. However, because of the greater relative strength of emission lines in the night sky and the non-uniform sky illumination during the twilight periods, "twilight flats" do not always give good results. In this case the sky background (between galaxies and stars) of the night sky itself has to be used for flatfielding. Since at faint magnitudes a large fraction of the dark sky is covered by astronomical objects, different night-sky images have to be collected and combined to produce a flat-field image which fully covers the FOV of a detector. In order to avoid the introduction of noise by the photon statistics of the flat-field images, an adequate number of individual flats has to be co-added to assure that the relative photon-noise error of the averaged flats becomes negligible with respect to the error due to the (unavoidable) error caused by the photon statistics of the sky background of the science frames. Co-adding many flats is possible since the effective QE changes only slowly with time. For very precise photometric measurements these time variations and various other higher-order effects have to be taken into account (e.g., [180]).

A particularly cumbersome systematic effect in CCD detectors are optical interference fringes produced by internal reflections in the different layers of these detectors (Fig. 4.7). Since red photons penetrate deeper into the photosensitive layers, fringing is most severe in the red spectral range. Moreover, the fringes are particularly strong for images obtained in monochromatic light (e.g., of emission line objects) or through narrow-band filters. Since flat-field images normally do not have

exactly the same wavelengths and the same incidence angles of the light, flatfielding normally cannot (fully) remove interference fringes. But, as shown by Fig. 4.7, the fringes can be eliminated by co-adding dithered images (see Sect. 4.2.2).

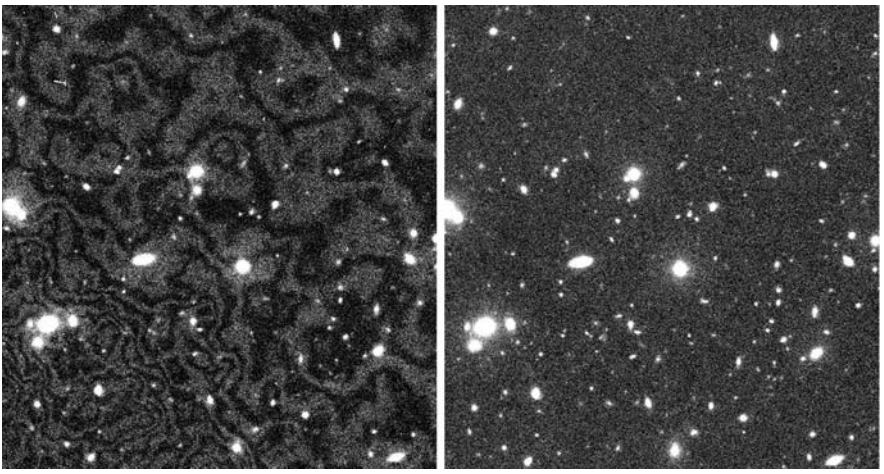

Fig. 4.7 Example of optical interference fringes occurring in CCD detectors. The two images show a galaxy field observed in the relatively narrow z-band filter (see Table 4.1). The *left panel* shows a raw image frame. The *right side* shows a reduced image based on co-added dithered individual frames. As shown by the figure, the reduction completely removed the fringe pattern (and other artifacts) of the raw image

Flatfielding spectral images is less complex since (apart from undispersed stray-light) the incident light at a given pixel is monochromatic. Therefore, for spectroscopic observations artificial continuum sources, such as incandescent light bulbs, can be used.

In addition to photons, solid-state detectors also record energetic particles due to cosmic rays and natural radioactivity. The background due to energetic particles can be reduced by avoiding materials containing radioactive isotopes in astronomical instruments. Since most glasses contain radioactive isotopes particular care has to be taken in the selection of the material of lenses and of the entrance window of detector cryostats. However, at telescope sites a complete exclusion of particle radiation is not feasible Therefore, astronomical images generally are contaminated by traces of energetic particles. In the astronomical literature these traces are referred to as "cosmics." Since the the photo-sensitive layers of astronomical detectors are rather thin, particles normally affect only a single pixel or a few adjacent pixels, which often are saturated. Therefore, the "cosmics" appear sharper than the PSF resulting from the optics and the seeing. A method of eliminating cosmics is fitting Gaussians to all apparent point sources detected in an image, followed by a removal of all apparent sources which have a FWHM below the PSF of the instrument.

After flatfielding and the removal of the cosmics, the sky background has to be determined (e.g., by evaluating histograms of the flux recorded by the individual

pixels) and subtracted from each individual frame. Since an incorrect background subtraction can lead to systematic errors of the photometry, any change of the background over the FOV has to be determined and taken into account(e.g., [228]). After the sky subtraction the individual frames taken for a given scientific image are co-added. As pointed out in Sect. 4.2.2 the individual frames are normally taken with small relative offsets. Therefore, before the co-adding, the frames have to be carefully recentered and corrected for possible differential image distortions. For each final image pixel the variations between the flux values obtained from the individual frames can be used to estimate statistical errors. Pixel values which show a statistically significant deviations (possibly caused by undetected cosmics or by detector defects) are eliminated from the final co-added images.

As final step, the images have to be calibrated by means of objects with known brightness in the same field or in a field observed with the same instrument under the same conditions. In principle, with linear detectors a single object with known flux is sufficient for this purpose. In practice several objects of different magnitudes should be observed to estimate the errors of the calibration and to test the linearity.

4.2.3.2 Extracting Source Positions, Radiation Fluxes, and Object Types

Images of the sky are normally taken to get information on individual objects in the observed field. If no prior information on the positions of the objects to be investigated is available, one first has to make an inventory of all recorded individual targets. Deep images obtained with modern large detectors at big telescopes typically record 10^4–10^5 individual objects. Evaluating such data sets by hand is impossible. Therefore, automatic "extraction programs" have been developed which determine the positions, fluxes, and some basic properties of all objects in a reduced image. A frequently used and publicly available code for detecting individual objects in ground-based deep optical images is the SExtractor code of Bertin and Arnouts [49], available, e.g., at `www.cfht.hawaii.edu/~morrison/home/SExtractor.html`. An object is regarded to be a real detection if a number of adjacent pixels (consistent with the optical PSF) show a flux of a pre-selected significance level. The number of adjacent pixels and the significance level are input parameters. Since marginally significant sources are of little value (as their properties cannot be determined) normally the parameters are chosen to give an overall significance of the detections of at least 5 σ or 10 σ. By fitting a suitable 2-D distribution function to the observed light distribution the shape, orientation, and the center position of the observed source can be determined. By comparison of the observed light distributions with the PSF (derived from stars in the field) point sources can be identified. By integrating the flux over the whole object or by integrating the flux inside a (software-defined) aperture, for each source the total flux or an "aperture flux" can be derived. While the total flux provides information on the object's luminosity, aperture flux values, which are better defined and more accurate, can be used to derive flux ratios and color indices. More sophisticated codes can determine the surface brightness distribution of (sufficiently) extended objects to provide an approximate morphological classification of the observed galaxies.

As with all automatic procedures, for each given image data the extraction codes have to be tested (e.g., using artificial data with known properties) and the dependence of the results on the input parameters has to be investigated. Among the pitfalls to be avoided is the dissection of low-surface brightness galaxies into several objects. Such a spurious splitting of a galaxy can occur if the flux in part of the object pixels is below the chosen significance limit, while other parts are above the limit. Space-based imaging is particularly prone to this effect, since (due to the low background) objects with a relatively low surface brightness can be detected, while (due to the relatively small apertures of space telescopes) the observed flux per pixel is fluctuating due to the object photon statistics. This effect can lead to incorrect number counts of the faint objects. In the case of ground-based observations this effect is less critical, since many more photons have to be collected for a significant detection against the brighter sky background.

Most critical in the case of ground-based images are spurious detections due to small local errors in the background subtraction. Therefore, the behavior and the results of extraction programs at the faint end have to be carefully tested.

Detailed discussions of the reduction and analysis of optical imaging data sets can be found, e.g., in [35, 221], and [228].

4.2.4 Calibrations, Units, and Sensitivities

4.2.4.1 Flux Units

Astronomy was one of the first sciences where quantitative measurements were made. Moreover, often special measuring techniques had to be developed which are not or only rarely used outside astronomy. Therefore, the astronomers initially developed their own systems measuring units. Today, following guidelines of the International Astronomical Union, astronomical results are presented using the international (SI) system of units, or units which are defined on the basis of the SI system. However, because of the historic development, many special terms and notations can still be found in the current astronomical literature. In the context of imaging and photometry this applies to some of the units used for measuring the radiation flux F and the spectral radiation flux F_ν or F_λ from astronomical objects.

In addition to the standard SI units (W m^{-2} for F, and W m^{-2} Hz^{-1} for F_ν), and closely related units (such as erg s^{-1} cm^{-2} Å^{-1} for F_λ) at present, essentially three other basic units are in regular use in the astronomical literature. The most frequently used astronomical flux unit is the "Jansky" (Jy)[3], which is defined as

$$1\,\mathrm{Jy} = 10^{-26}\,\mathrm{W\,m^{-2}Hz^{-1}} = 10^{-23}\,\mathrm{erg\,s^{-1}cm^{-2}Hz^{-1}}. \tag{4.4}$$

[3] Named after the pioneer of radioastronomy Karl G. Jansky

The unit Jy was initially introduced in radio astronomy at a time when 1 Jy was approximately the sensitivity limit of radio telescopes. However, since this unit is a convenient measure for the small fluxes typically encountered in astronomy, it is now used throughout the electromagnetic spectrum.

Another measure for radiation fluxes (and spectral fluxes) is the "photon flux," defined as the number of photons recorded per unit of time and per m^{-2} (or photons recorded per unit of time per m^{-2} and per unit photon energy, or per Hz). Thus photon fluxes typically are given in units of $s^{-1}m^{-2}$, $s^{-1}m^{-2}$ eV^{-1}, or m^{-2}. These units are particularly popular in X-ray, FUV, and Gamma astronomy. Obviously, "photon fluxes" can be readily converted into SI units, if the photon energy distribution and the instrumental efficiency as a function of photon energy are known.

The traditional units for optical and infrared radiation fluxes are the "magnitudes" m. Magnitudes are dimensionless, logarithmic measures of the radiation flux F in a selected wavelength band (usually defined by a filter transmission curve) according to

$$m = -2.5 \log \frac{F}{F_0} = -2.5 \log F + 2.5 \log F_0 \tag{4.5}$$

where F is the observed Flux in the filter band in question, and F_0 is the flux observed in the same filter band in a reference spectrum, which defines the zero point of the system.

Magnitudes (and other astronomical flux data) always give the flux reaching the Earth outside the Earth's atmosphere. Thus ground-based observations have to be corrected for losses due to absorption and scattering in the atmosphere. These losses can be determined by means of flux measurements at different zenith distances (corresponding to different pathlengths through the atmosphere). The flux outside the atmosphere is then derived by extrapolating to a zero pathlength.

Currently, two different magnitude systems are in popular use, the "AB" magnitudes and the "Vega" magnitudes. The two magnitude systems differ with respect to the functions F_0. The AB system [383, 384] assumes for F_0 in Eq. (4.5) a constant spectrum

$$F_\nu = 3.63 \times 10^{-23}\,\mathrm{W\,m^{-2}Hz^{-1}}. \tag{4.6}$$

Thus, for an object from which we receive (outside the Earth's atmosphere) a constant spectral flux $F_\nu = 3.63 \times 10^{-23}\,\mathrm{W\,m^{-2}Hz^{-1}}$ we have at all wavelengths and for all filter transmission functions AB magnitudes $m = 0$. Correspondingly, objects with the same spectrum which are fainter by a factor 0.1 we have (according to Eq. (4.5)) AB magnitudes of $m = 2.5$, and objects with a constant F_ν = 1 Jy have $m = 8.9$. If F_ν is not constant, the magnitude depends on the central wavelength and on the shape of the filter function.

The "Vega" system [253] uses as reference spectrum the observed energy distribution F_ν of the bright star Vega (α Lyrae). Between about 380 nm and 548 nm the spectral flux observed from Vega is slightly higher than the AB reference flux

(Eq. 4.6). Therefore, in this wavelength range the numerical values of the Vega magnitudes are slightly (up to 0.3) higher numbers. At shorter and longer wavelengths, the Vega spectrum falls below the AB reference spectrum. Therefore, at these wavelengths the numerical values of the AB magnitudes are larger numbers. For an object with a flat F_ν spectrum the "visual" magnitude m_V has the same value in the AB and the Vega system. For the filter bands U, R, and I the differences $m_{AB} - m_{Vega}$ are approximately 0.9, 0.2, and 0.4, respectively. The exact differences depend on the spectrum of the object and on the filter function.

Vega magnitudes obviously have the advantage that photometric instruments can be directly calibrated by observing Vega (or a star with a similar spectrum which has been calibrated by means of Vega). On the other hand, conversion of Vega magnitudes into SI units requires an accurate knowledge of the Vega spectrum. Therefore, systematic errors in the Vega spectrum can cause erroneous photometric results.

As pointed out above, the wavelength band corresponding to the flux F in Eq. (4.5) normally is defined by a filter function. For historic as well as for astrophysical reasons many different filter systems have been developed and are being used. The different systems differ with respect to the central wavelengths, the widths, and the peak transmission of the individual filters. Apart from narrow-band filters, designed to measure the flux in selected spectral lines, there exist intermediate-band filters (with widths of the order 10 nm) and broad-band filters (with widths of the order 100 nm). Among the most frequently used broad-band systems are the Johnson–Cousins–Bessell system (UBVRI, see [51, 52, 253]), the ugriz system of the Sloan Digital Sky Survey [482], and the F363W – F814W system of the WFPC2 camera on the Hubble Space Telescope (e.g., [233]). The basic properties of these three systems are listed in Table 4.1. Figure 4.8 gives the transmission functions of the U, B, g, R, and I filters of the FORS cameras of the ESO VLT.

Table 4.1 Effective wavelengths λ_{eff} (in nm) and FWHM band widths $\Delta\lambda$ (in nm) of three frequently used broad-band photometric systems

Filter	λ_{eff}	$\Delta\lambda$	Filter	λ_{eff}	$\Delta\lambda$	Filter	λ_{eff}	$\Delta\lambda$
U	366.3	65	u'	359.6	57	F336W	344.8	34
B	436.1	89	g'	463.9	128	F439W	430.0	72
V	544.8	84	r'	612.2	115	F555W	532.3	155
R	640.7	158	i'	743.9	123	F675W	666.7	123
I	798.0	154	z'	889.6	107	F814W	787.2	146

At wavelengths $>0.9\,\mu$m ground-based astronomical photometry is severely hampered by molecular absorption lines of water vapor and other atmospheric molecules. Since the water content of the atmosphere depends on the temperature and geographic location, the H_2O absorption (and emission) varies strongly in space and time. Therefore, ground-based IR-observations try to avoid the H_2O absorption by using filter bands which coincide with gaps in the H_2O spectrum. Examples of such H_2O-avoiding NIR filter bands are the Y ($\lambda_{eff} = 1.0\,\mu$m), J ($\lambda_{eff} \approx 1.25\,\mu$m), H ($\lambda_{eff} \approx 1.63\,\mu$m), and K ($\lambda_{eff} \approx 2.2\,\mu$m) bands. The exact effective wavelengths of the NIR bands differ between the individual IR instruments. Minor

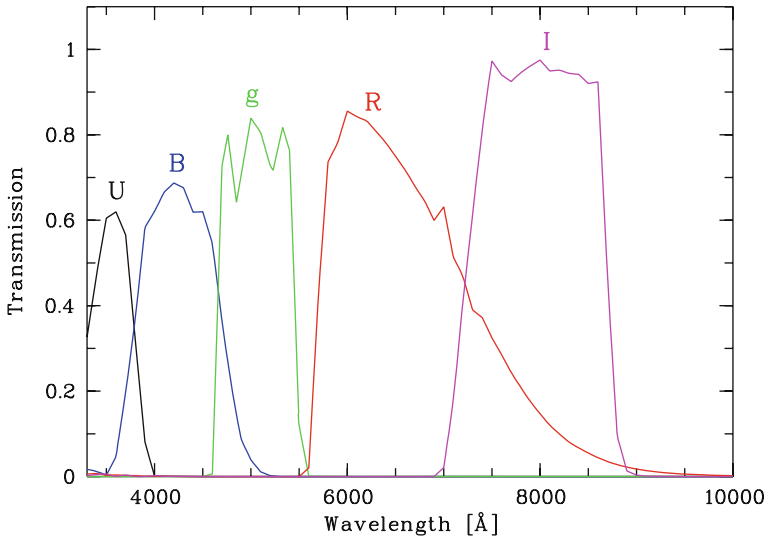

Fig. 4.8 Examples of optical filter transmission curves. Shown are the transmission functions of the U, B, g, R, and I filters of the FORS cameras at the ESO VLT

modifications of the IR filter bands are sometimes indicated by a subscript to the band designation. An example is the (frequently used) K_s-band, which differs from the standard K-band by a slightly shorter effective wavelength and a narrower band widths (which avoids sky-background lines). A comprehensive discussion of current optical and IR photometric filter systems and graphic presentations of the corresponding response functions can be found in [52].

To specify the filter band for which a magnitude has been measured, m with a corresponding subscript (e.g., m_V or m_{F555}), or the designation of the filter band itself is used instead of m. Thus, either m_V or V can stand for the radiation flux (expressed in magnitudes) in the V ("visual") band of the UBVRI system.

Magnitudes are also used to measure flux ratios, which according to Eq. (4.5) correspond to magnitude differences. In the astronomical literature broad-band magnitude differences (such as $B - V$) are called "color indices" (or simply "colors"). As pointed out in Sect. 2.1.2, color indices can be used as a measure of the color or the effective temperature of a star. Plots relating the flux (expressed in magnitudes) of stars or galaxies with flux ratios (in magnitudes) are called "color-magnitude diagrams." An important example of a color-magnitude diagram is the Hertzsprung–Russell diagram introduced in Sect. 2.1.2. Correspondingly, plots relating two different flux ratios are called "color-color diagrams."

In addition to measuring fluxes in wavelength intervals defined by filter bands, magnitudes are also used to describe the total flux

$$F_{total} = \int_0^\infty F_\nu \mathrm{d}\nu. \tag{4.7}$$

received from an object. The corresponding "bolometric magnitude" m_{bol} is defined as

$$m_{bol} = -2.5 \log \frac{F_{total}}{F_{t0}} \tag{4.8}$$

where (in the Vega system) $F_{t0} = 2.52 \times 10^{-8}\,\mathrm{W\,m^{-2}}$.

Finally, in addition to the normal magnitudes m, which measure the flux received by us from astronomical objects, there are the so-called "absolute magnitudes" M which are a measure of the luminosity of the sources. M is defined as the "apparent" magnitude m which an object would have at a distance of 10 pc. For small distances ($z \ll 1$) we have

$$M = m + 5 - 5 \log D \tag{4.9}$$

where D is the distance in parsec. The difference $m - M$ is called "distance modulus." It obviously provides a logarithmic measure of the distance.

4.2.4.2 Flux Calibration

In laboratory physics photometric instruments are calibrated by measuring the flux from a radiation source of accurately known luminosity and spectrum. Suitable radiation sources are, e.g., a blackbody in thermodynamic equilibrium or a synchrotron radiation source. Observing such laboratory sources with large optical telescopes is not feasible. Therefore, astronomical photometric measurements normally are calibrated using bright standard sources on the sky, such as stars or planets. These astronomical standard sources are calibrated against laboratory standard radiation sources using small telescopes which can observe the sources on the sky as well as the laboratory radiation sources (or reliable intermediate standards, accurately calibrated in the laboratory). However, an astronomical object normally can be observed only during part of the year and only from part of the Earth's surface. Moreover the primary standards are too bright to be observed with large telescopes. Therefore, usually systems of secondary standard sources are defined and calibrated by comparison with the primary standards.

The primary standard for the optical spectral range is the star Vega. The flux and spectrum of Vega has been carefully calibrated independently by several different groups.

In principle an astronomical instrument can also be calibrated by measuring the absolute sensitivity of the detector or receiver in the laboratory and by determining the efficiencies of all optical components involved. In practice this does not give accurate results since the exact efficiency of large optical components (such as large telescope mirrors) are difficult to measure and since the errors of the measurements of many individual components tend to add up to significant uncertainties. A comparison of absolute flux values based on the Vega calibration with values using calibrated detectors and calculated optical efficiencies often results in differences

which are significantly larger than the calculated observational errors. This may indicate that the errors in the presently used calibrations are larger than given in the literature.

Absolute flux calibrations in the radio range are simpler, since here the standard sources can be directly compared with low-temperature blackbodies and other reliable standard sources. In the FUV and EUV range the calibration rests on theoretical models of the atmospheres of hot white dwarf stars. Since the structure of these objects is relatively well understood, the corresponding calibration are regarded as reliable. The relatively small X-ray telescopes can be directly calibrated using laboratory X-ray sources.

4.2.4.3 Photometric Accuracy and Sensitivity Limits

As pointed out in Sect. 4.1 the accuracy of photometric measurements is invariably limited by the photon statistics. If n photons are recorded, the photon statistics results in an absolute mean error of $\sqrt{n}$, a relative mean error of $1/\sqrt{n}$, and a signal-to-noise-ratio (S/N) of $\sqrt{n}$. The actual S/N of photometric observations is lower, since there always exist additional noise sources, which increase the denominator. For the important case of CCD photometry the S/N for each pixel normally can be approximated by the relation

$$S/N = \frac{\Delta t\ P_{object}}{\sqrt{\Delta t\ P_{object} + \Delta t\ P_{sky} + \Delta t\ dc + RON^2 + \Delta t\ P_{sky}\ \sigma_{sky}}} \tag{4.10}$$

where Δt is the integration time, P_{object} the number of photons recorded per second from the observed object, P_{sky} the number of photons recorded from the sky background, dc the dark current, RON the readout noise, and σ_{sky} is the rms error of the sky background derivation. If the dark current and the RON are negligible relative to the other terms (which normally is the case for faint objects), Eq. (4.10) can also be applied directly to extended targets, such as galaxies. (In this case for P_{sky} the background integrated over the surface of the object has to be used in Eq. 4.10.)

From Eq. (4.10) it is obvious that for bright objects ($P_{sky} \ll P_{object}$) the integration time Δt required to reach a given S/N increases $\sim P_{object}^{-1}$. If $P_{object} \ll P_{sky}$ and $\Delta t\ P_{sky}\ \sigma_{sky} \ll \sqrt{\Delta t\ P_{sky}}$ the integration time increases $\sim P_{object}^{-2}$. If $\Delta t\ P_{sky}\ \sigma_{sky}$ becomes the dominant term in the denominator of Eq. (4.10), the S/N becomes independent of Δt. This again confirms the critical importance of an accurate background derivation for the photometry of faint objects.

For the V-band of the UBVRI photometric system the spectral photon flux (expressed in photons $s^{-1}\ m^{-2}\ Å^{-1}$) is approximately

$$P(m_V) \approx 10^{7-0.4\,m_V}. \tag{4.11}$$

High-redshift galaxies typically have magnitudes fainter than $m_V = 20$. According to Eq. (4.11) at this magnitude (corresponding to about 36 μJy) the spectral photon flux is about 0.1 photons $s^{-1}\ m^{-2}\ Å^{-1}$. If we observe a $m_V = 20$ galaxy with a

8-m telescope (corresponding to a collecting surface of $\approx$ 50 m^2) we get about 5 photons s^{-1} Å^{-1}. At $m_V = 25$ (0.4 μJy) we get 0.05 photons s^{-1} Å^{-1}, and at $m_V = 30$ (4 nJy) about 2 photons hour^{-1} Å^{-1}.

In order to estimate the S/N from Eq. (4.10) in addition to the target flux the sky background of the area covered by the target (or, in the case of point sources, the area of the seeing pattern) has to be known. For ground-based observations in a moonless night the sky background in the V-band is of the order $m_V = 21.6$ per arcsec^2. Thus, from a circular sky area of one arcsec diameter, an 8-m telescope collects about one V-band photon s^{-1} Å^{-1}. For a 0.5 arcsec diameter area this value drops to about 0.2 photons s^{-1} Å^{-1}. As shown by these numbers, in the case of ground-based observations the flux observed from high-redshift galaxies is normally below that of the underlying sky background, and the photon noise of the sky background normally is the dominating noise source. Good seeing conditions and a high surface brightness of the target obviously improve the S/N.

Most of the ground-based sky background originates from line emission of molecules and atoms in the Earth's atmosphere and from scattered continua from various sources. Therefore, the sky background varies strongly with wavelengths (Fig. 4.9). In order to reach high sensitivities, photometric wavelength bands have to be selected which avoid strong sky emission features. Lists of strong night-sky emission lines can be found in astronomical and geophysical handbooks, and in some publication dedicated to this topic (e.g., [391]). In the case of space-based observations the main sources of the optical and IR sky background is solar radiation scattered by interplanetary dust particles, atoms and ions. Apart from resonant

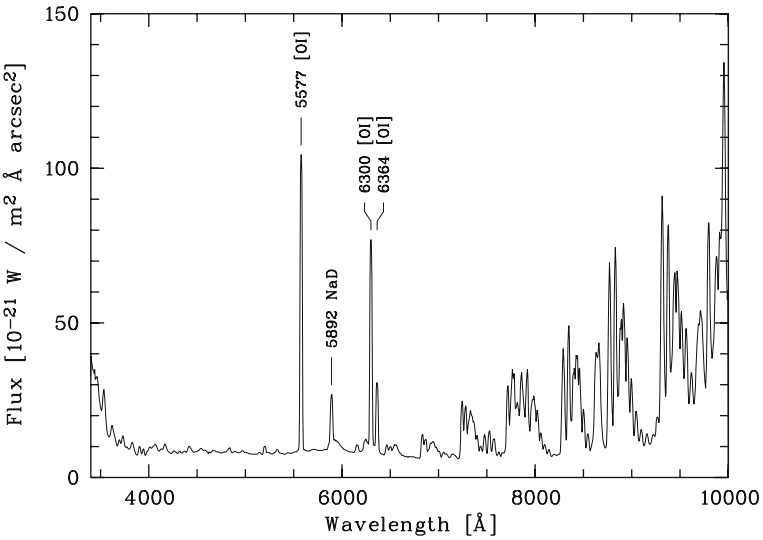

Fig. 4.9 Spectrum of the (moonless) night sky at the ESO Paranal observatory in Northern Chile. (Courtesy S. Noll)

scattered emission lines (such as Lyα), the spectrum of the background outside the Earth atmosphere is essentially a solar spectrum.

At low radio frequencies the synchrotron emission of the interstellar medium dominates the background. At high radio frequencies, the sensitivity normally is limited by the thermal emission from the Earth's atmosphere or by the receiver noise. (For details on the sensitivities in the radio range see [441].)

In order to optimize observing programs and to avoid wasting of valuable observing time, astronomical observations are normally simulated using computer models and empirical relations such as Eq. (4.10) before going to the telescope. Corresponding "exposure time calculators" (ETCs) can be accessed through the web pages of the corresponding observatories. With such programs for many types of observations fairly reliable S/N values and sensitivity limits can be predicted.

4.3 Spectroscopy

The most important source of information on the physical conditions and processes in the distant universe is the analysis of astronomical spectra. Particularly productive have been observations of the line spectra of atoms, ions, and molecules which can provide accurate information on the composition, temperature, density, and (via the Doppler effect) the motions of the observed matter. Apart from the fact that with increasing distance the flux becomes lower and the spectra are redshifted to different wavelengths, spectral analysis can be carried out completely independent of the distance of the source. Thus, spectroscopy is an almost ideal remote sensing technique, even on cosmic scales.

The information content of spectra depends on the S/N and on the spectral resolution. In the modern astronomical literature the spectral resolution is defined as $R = \lambda/\Delta\lambda$, where λ is the observed wavelength and $\Delta\lambda$ is the observed FWHM width of a line with an intrinsically negligible width. With increasing R more detailed line profiles and weaker lines can be observed. Thus, in principle a high spectral resolution is always of advantage. However, from Eq. (4.11) and its discussion it is clear that at a given total flux with increasing spectral resolution the number of photons (and consequently the S/N) per resolution element is decreasing. As a result, for faint objects the maximal attainable spectral resolution is about proportional to the observed flux.

4.3.1 Spectrographs

Today spectroscopic astronomical observations are being carried out at all wavelengths of the electromagnetic spectrum. However, in different spectral regions different techniques are used. At Gamma and X-ray wavelengths usually the energy of the individual photons is measured directly during the detection. At X-ray wavelengths this can be done using gas-ionization or solid-state detectors with band gaps

much smaller than the observed photon energies. Such detectors produce for each detected photon a charge signal which is proportional to the photon energy $h\,\nu$.

Using superconducting junctions this technique can also be applied to optical wavelengths. However, at present, large-format superconducting junction arrays are not yet available. Therefore, for optical and IR spectroscopy the photons are normally detected without energy information. Instead, an optical image of the spectrum is produced and recorded by a normal imaging detector (of the types described in Sect. 4.2.2). In most cases the spectra are produced using diffraction gratings or (less often) prisms. Modern blazed gratings and grating-prism combinations ("grisms") have high efficiencies and can be made for a large range of wavelengths, spectral resolutions, and transmission characteristics.

In principle spectra of stars and galaxies can be obtained by placing a transmission grating or a prism in front of the entrance aperture of a telescope. In this case for each object (instead of an image) a spectrum is produced in the focal plane. Obviously, this method (called "objective grating" or "objective prism" technique) cannot be directly applied to very large telescopes. However, the method can be adopted to telescopes of any size by placing a transmission grating, prism, or grism into the parallel beam between the collimator and the camera of a focal reducer.

A disadvantage of the objective grating technique is that the spectra of different objects in the field can overlap and that the full flux from the sky is imaged onto the detector. Because of the resulting high background, the objective grating technique normally cannot be used for faint objects, such as high-redshift galaxies. A notable exception are objective grating observations of high-redshift emission-line objects which have been successfully carried out using the Hubble Space Telescope ([402]), making use of the low background in space.

Overlapping spectra and a high background can be avoided by placing a (or many) spectrograph slit(s) into the focal plane. In this case only the light entering the slit(s) contributes to the background of a spectrum. Instead of slits the polished ends of optical fibers can be placed into the focal plane. If the light exiting the output ends of the fibers is fed into a spectrograph, for each fiber a spectrum is recorded. An example of a large fiber-feed assembly is shown in Figs. 4.10 and 4.11. With suitable slit or fiber-feed assemblies up to several hundred spectra can be obtained simultaneously with such "Multi-Object Spectrographs" (or "MOS" devices). For very faint objects the multi-slit technique normally is the better choice, since the background suppression can be optimized by adjusting the slit width, and since light losses by scattering and absorption in the fiber material are avoided. The fiber technique is superior for observing objects distributed over large fields which cannot be well imaged directly through a spectrograph.

Tightly packed fiber ends or similar optical systems also can be used to dissect an image into many adjacent image elements for which individual spectra are produced. This technique, called "integral field spectroscopy," is particularly useful to study the spectral properties and physical parameters over the surface of a galaxy. Current integral field spectroscopy units ("IFUs") produce up to about 10^3 individual spectra.

Fig. 4.10 Fiber-feed assembly for the fiber-fed multi-object spectrograph at the AAO 3.9-m telescope. This instrument was used for the 2dF Galaxy Redshift Survey ([106]). The assembly consists of 400 fiber ends on top of a metal plate in the focal plane of the telescope. Before each exposure the fiber ends are positioned at the coordinates of the spectroscopic targets by computer-controlled robot. During the exposure the feeds are magnetically attached to the focal-plane plate. (Photo and copyright: Anglo-Australian Observatory)

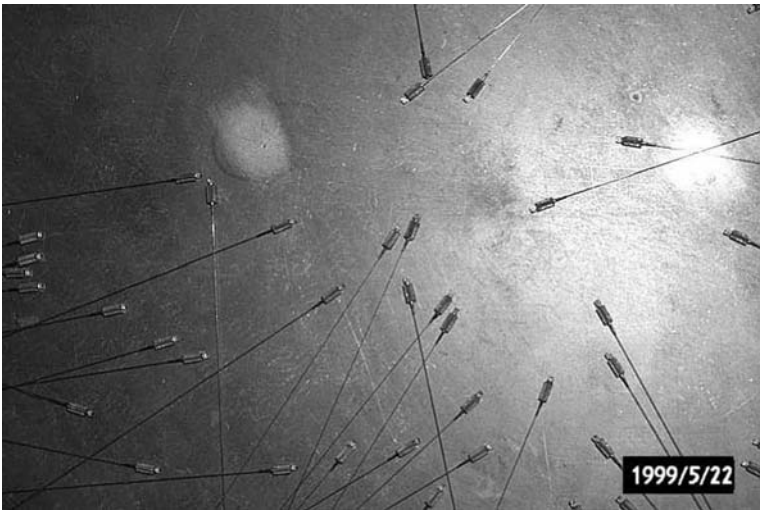

Fig. 4.11 Close-up of the fiber ends of the 2dF fiber-feed assembly. At each fiber end a prism reflects the light (coming from above) into the corresponding fiber. Just behind the prisms are the small magnets holding the fiber ends in position.(Photo and copyright: Anglo-Australian Observatory)

In addition to grating and grism spectrographs sometimes Fabry-Perot (FP) spectrometers are used in astronomy. The main application of FP systems is imaging in very narrow spectral bands. Due to the high spectral resolution of high-order FP systems, quasi-monochromatic images can be obtained.

For spectroscopic observations at long wavelengths (sub-millimeter, millimeter, and radio range) FP-like techniques, narrow-band receivers, electronic filter banks, acousto-optical spectrometers, and various Fourier techniques are used. For details on these methods see, e.g., [441].

Detailed technical information and descriptions of the spectrometers which are currently available at the major astronomical observatories can be found on the web pages of the corresponding institutions.

4.3.2 Deriving Intrinsic Luminosities and Spectra

Astronomical spectrographs record the spectra as they appear in the rest-frame of the observer. Due to the expansion of the universe these spectra differ from the those emitted by the observed target. Since the interpretation of the spectra in terms of the physical conditions in the target requires the knowledge of the intrinsic spectral properties, the observed data have to be converted to intrinsic spectra. Apart from the wavelengths, which can be readily converted according to Eq. (1.2), most important is the conversion of the spectral fluxes F_ν and F_λ to intrinsic spectral luminosities.

A relation connecting the observed total radiation flux of a distant object with its emitted luminosity has been given in Eq. (3.25). To derive a relation between the observed spectral flux density F_ν or F_λ and the intrinsic spectral power L_ν, one has to take into account that the wavelengths are shifted and that the power which is emitted in a given intrinsic wavelength interval $\Delta\lambda$ is redistributed into a redshifted broader wavelength interval $\Delta\lambda(1+z)$. By adding these two effects to Eq. (3.25), we get for the observed spectral flux:

$$F_\nu = \frac{L_{\nu e}}{4\pi a_0^2 r^2 (1+z)} \tag{4.12}$$

and

$$F_\lambda = \frac{L_{\lambda e}}{4\pi a_0^2 r^2 (1+z)^3}. \tag{4.13}$$

where ν and λ are the observed frequency and wavelength, $L_{\nu e}$ and $L_{\lambda e}$ denote the spectral power emitted at the frequency $\nu(1+z)$ and wavelength $\lambda(1+z)^{-1}$, respectively.

By inverting these relations the intrinsic spectra are derived from the observed redshifted data.

The shift of the effective wavelength and the broadening of spectral bands obviously also affects photometric measurements of redshifted objects. Using Eq. (4.12)

or (4.13) the relation between observed and intrinsic fluxes can be calculated for the various photometric systems. The effect of the different distance dependence of observed filter-magnitudes and of bolometric magnitudes is often expressed by the so-called "K-correction" K, which is defined as $K = m - M - (m_{bol} - M_{bol})$, where the subscript *bol* denotes bolometric magnitudes. In order to calculate K, the redshift, the filter characteristics, and the intrinsic spectrum have to be known.

In order to get accurate intrinsic spectral fluxes and luminosities from the relations given above, an additional correction has to be applied to remove the flux attenuation by the galactic interstellar medium. This attenuation is mainly due to interstellar dust. Therefore, it is wavelength dependent (e.g., [287]) and affects the total flux as well as the flux ratios or color indices. Since the total galactic dust absorption of extragalactic objects cannot be directly measured, it is normally inferred from the observed galactic neutral hydrogen column density in the corresponding direction. Maps of the resulting optical dust absorption as a function of the position on the sky can be found in [82].

For wavelengths shortward of the Lyα line, luminosities of distant galaxies are also attenuated by H I absorption in the intergalactic medium (see Sect. 6.9). A reliable correction of this effect requires high-resolution spectra.

A spectroscopic quantity which is also affected by the cosmic expansion is the "equivalent width" W_λ of a spectral line. W_λ originally has been defined for absorption lines by

$$W_\lambda = \int \frac{F_c - F_\lambda}{F_c} \mathrm{d}\lambda \tag{4.14}$$

where F_c is the (interpolated) continuum level at the position of the absorption line. W_λ has the dimension of a wavelength (and usually is given in Å), but it is a measure of the strength of a line. For an absorption line with a box profile with zero rest intensity W_λ becomes equal to the line width. The definition of W_λ (Eq. 4.14) obviously can be generalized to include emission lines, for which Eq. (4.14) results in $W_\lambda < 0$. "Emission line equivalent widths" are defined by reversing the sign of Eq. (4.14).

For stellar spectra equivalent widths can be predicted accurately from model atmospheres. By combining stellar models and photoionization models, equivalent widths can also be predicted for the H II regions of galaxies. If $z \ll 1$ the equivalent width W_λ is independent of the distance of an object. Therefore equivalent widths are a particularly useful line strength measure for astrophysics. However, since the redshift stretches the wavelength scale by a factor $(1+z)$, while the normalization to the continuum remains unchanged, for a redshifted object the observed equivalent width W_λ becomes

$$W_\lambda = W_{\lambda\,0}(1 + z) \tag{4.15}$$

where $W_{\lambda 0}$ is the equivalent width in the intrinsic (emitted) spectrum. Hence observed equivalent widths of high-redshift objects have to be corrected according to Eq. (4.15).

4.4 Virtual Observatories and Archives

Most astronomical images and spectra are taken in the context of a specific scientific program. They are obtained to derive an answer to a specific scientific question. But quite often images and spectra contain, in addition, information which can be useful or vital for other astrophysical problems. To make the reuse of astronomical images and spectra for such other applications possible, astronomical data are traditionally put into archives and data banks which are accessible to the astronomical community. Astronomical data collections in the form of catalogs predate the Ancient Greek astronomy, and have been accumulating ever since. After the invention of the photographic process, plate collections of astronomical images and spectra became the most important archives. Today, at all major astronomical observatories the observed data are archived in digital form, and much of the old photographic-plate data have by now also been scanned and digitized. These digital data bases, which are rapidly growing, have become an important tool for modern astronomy.

Astronomical data bases can consist of an archive of raw observational data, or of reduced observations, or the data may be accessible in the same form and with the same auxiliary data and reduction software as provided by modern observing facilities. In the latter case the data base forms a true "virtual observatory" which can be used like a real astronomical observing facility. However, the access to virtual observatories is much faster and usually easier than to the heavily oversubscribed modern large telescopes.

Astronomical data archives and virtual observatories are accessed through their web sites. Among the most important astronomical data bases is the "Centre de Données astronomiques de Strasbourg" (CDS) (http://cdsweb.u-strasbg.fr) which also provides links to all other major astronomical archives. Links to most existing or nascent virtual observatories can be found on the web page of the "International Virtual Observatory Alliance" (IVOA, *www.ivoa.net*) which (at present) represent 16 VOs in different countries. A fairly complete data base for the scientific astronomical literature is the "Smithsonian/NASA Astrophysics Data System" (ADS, http://www.adsabs.harvard.edu). An important source of infrared astronomical data is the web site of the "NASA Infrared Processing and Analysis Center" (IPEC, http://www.ipac.caltech.edu). Among the other archives operated by NASA are optical data collections at the (Hubble) Space Telescope Science Institute (STScI, http://www.stsci.edu). An archive for astronomical software can be accessed through www.skysoft.org.

Chapter 5
Finding Very Distant Galaxies

I-band images recording all galaxies to a limiting magnitude of $m_I = 28$ contain about 10^6 galaxies per square degree. About 75% of these galaxies are foreground objects with redshifts $z < 2$, and only about 3% have redshifts $z > 4$. Thus, search procedures aiming at the identification of high-redshift galaxies must be capable of handling large data sets. Moreover, these procedures must reliably separate the high-redshift objects from the more abundant foreground objects. According to the numbers mentioned above, a potential misidentification rate as low as a few per cent can severely contaminate samples of high-redshift objects.

5.1 Identification Techniques for High-Redshift Objects

5.1.1 Spectroscopic Methods

The most direct and most reliable procedure for finding high-redshift galaxies in a given field is taking spectra of all detectable objects. Redshifts can then be determined directly by measuring the wavelength of spectral lines, or by a comparison of the observed spectra with template spectra of known redshift. In principle, such a complete spectroscopic coverage of a sky area is possible by means of integral-field spectroscopy techniques. However, at present, the modest sizes of the available IFUs limit such observations to relatively small fields.

Instead of IFUs, in principle, the "objective grism" technique (cf. Sect. 4.3) can be used to get a complete spectroscopic coverage of a field of view. Currently available objective grism instruments cover larger fields than the present generation of IFUs. Like in the case of integral-field spectroscopy, with the objective grism technique the spectra of all objects in the field can be observed simultaneously. However, as explained in Sect. 4.3, objective grism instruments have the severe disadvantage of a high sky background and a correspondingly low signal-to-noise ratio. Therefore, apart from successful searches for emission-line objects with the HST (Sect. 5.1.3), the objective-prism method has, so far, not contributed much to the identification of high-redshift galaxies.

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_5,

5.1.2 Photometric Techniques

A standard method for identifying high-redshift objects in extended fields is the comparison of the observed flux ratios in different filter bands with those computed for galaxies of different redshifts. As discussed in Sect. 6.2.1, in this way reasonably accurate "photometric redshifts" can be obtained from images obtained in different filter bands. The accuracy and reliability of photometric redshifts depends on the number and of the widths of the filter bands used. Approximate redshifts can be obtained with a modest amount of effort, but accurate photometric redshifts require many bands and narrow filters, which results in long exposure times and significant amounts of observing time.

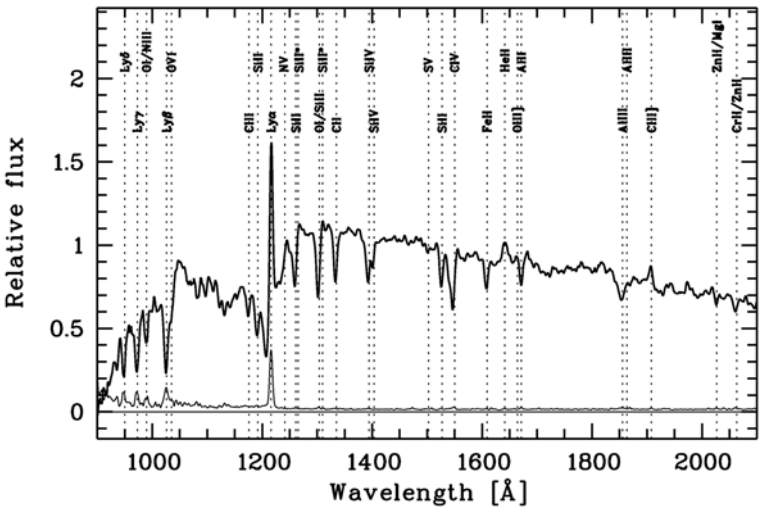

Fig. 5.1 Typical rest-frame UV spectrum of $z \approx 3$ star-forming galaxies. This spectrum has been produced by averaging the individual spectra of 64 galaxies with redshifts $2 < z < 4$. Note the drop of the continuum flux shortward of ≈ 1000 Å and ≈ 1200 Å. (From [377])

A related method is the "Lyman-break technique." This technique makes use of steps or breaks in the observed continua of star-forming high-redshift galaxies. The two most important breaks in the spectra of high-redshift galaxies are clearly visible in Fig. 5.1. The first of these breaks occurs at about 912 Å. This break is due to the Lyman-continuum absorption in the stellar atmospheres, in the (intrinsic) interstellar medium, and by the intergalactic atomic hydrogen shortward of 912 Å. As shown by Fig. 5.1, due to the Lyman continuum absorption, the spectral flux drops to almost zero for rest-frame wavelengths $\lambda < 912$ Å. A second pronounced step in the spectral flux occurs near 1215 Å. This step (again clearly visible in Fig. 5.1) is due to intergalactic Lyα line absorption. Since foreground Lyα absorbers can have any redshift between the target's redshift and $z = 0$, the IGM Lyα absorption can reduce the spectral flux at all wavelengths shortward of the Lyα line. As discussed in

Sect. 6.9, the density of Lyα absorbers increases with increasing redshift. Therefore, the amplitude and the wavelength of this second break depend on the redshift.

For redshifts about $z > 2$ the breaks in the SEDs of star-forming galaxies are shifted into the spectral range observable with ground-based optical telescopes. Since the breaks in the SEDs cause characteristic flux ratios (or characteristic color indices) in broad-band photometric measurements, optical broadband photometry can be used to identify the high-redshift objects.

The Lyman-break technique and the photometric redshifts both use flux ratios. But the two methods differ in the way these flux ratios are evaluated. For photometric redshifts the observed flux ratios are compared to template spectra in order to derive individual redshifts. The Lyman-break technique identifies galaxies of a certain redshift range directly from the flux ratios or color indices, making use of the fact that the high-redshift objects occupy distinct areas in suitably selected color–color diagrams. An example of such a diagnostic diagram is given in Fig. 5.2. A detailed discussion of the Lyman-break method and the selection criteria for different filter systems can be found in [203].

High-redshift galaxies which have been identified by means of the Lyman-break method are called "Lyman-break galaxies." Since a large fraction of the known high-redshift galaxies has been found using this technique, in the literature the term "Lyman-break galaxy" is often used synonymous with UV-bright high-redshift

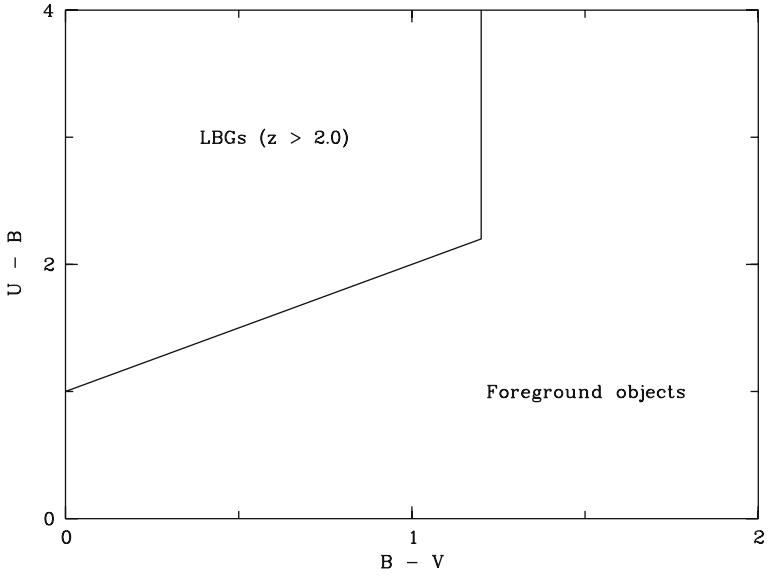

Fig. 5.2 Example of a diagnostic diagram for identifying high-redshift galaxies with the Lyman-break technique using HST data. The figure shows the location of $z > 2$ galaxies in a color–color diagram based on the U_{300}, B_{450}, and V_{606} filter bands of the HST WFPC2 camera. (According to [131] and [203])

galaxies, or for high-redshift galaxies in general. To minimize confusion, in this book the term "Lyman-break galaxy" is used only in the original sense.

Since the position of an object in the diagnostic diagrams of the Lyman-break technique can be derived from fewer and broader filter bands, this method requires less observing time than photometric redshifts. On the other hand, the original Lyman-break technique works only for star-forming galaxies and (for a given set of filters) for a certain redshift interval. Most early Lyman-break programs used U, B, V, (I) or U, G (= green), R, I filter sets, which are efficient for finding starburst galaxies in the redshift range $2.5 < z < 3.5$ (where the Lyman-continuum break moves through the U-band). Using different filter sets, the Lyman-break method later has been extended to objects with lower and higher redshifts (e.g., [5, 496, 572]). Using instead of U–B and B–V color–color diagrams diagnostic diagrams involving B–R and R–i, star-forming galaxies at $z \approx 4$ can be identified, while the use of the R, i, and z bands yield $z \approx 5$ objects.

As noted above, the Lyman-break technique does not provide redshifts of individual galaxies. But it allows us to estimate a mean redshift and a redshift range of galaxy samples. Moreover, by using spectroscopic subsamples or model computations, normally the redshift distribution of a given sample of Lyman-break galaxies can be determined. Examples of such distributions are presented in Fig. 5.3. If the distributions are well enough determined, samples of Lyman-break galaxies can be used efficiently for statistical studies, in spite of the fact that the redshifts of the individual sources are not known.

Historically, the Lyman-break technique has been the first method to yield major samples of high-redshift galaxies. It was first applied by Guhathakurta et al. in 1990 [214]. Later this technique was used extensively and with particular success by C. Steidel and his coworkers (e.g., [498, 500, 502]).

Because of the strong Lyman-continuum break, for galaxies with $z > 3.5$ normally no flux is observed in the U-band. Objects not visible on U-band images (but detected at longer wavelengths) are often referred to as "U-dropouts." Correspondingly, "B-dropouts" (with no flux at U and B), V-dropouts (no flux at U, B, and V), etc. are defined. While many U-dropouts are galaxies with redshifts $z > 3.5$, there exist other objects without measurable U-band flux (including moderately redshifted E-galaxies and red galactic stars). Therefore, a "dropout" in the U-band (or at any other filter band) alone cannot be used to constrain the redshift. However, additional flux ratios from other broadband filters or spectroscopic observations often make it possible to unambiguously determine the nature of a dropout object and to constrain its redshift and the redshift distribution of dropout objects. In this case dropout samples can also be used for statistical purposes. (Fig. 5.3).

If IR filter bands are included, color-index selection criteria can also be extended to identify passively evolving high-redshift galaxies. As shown by E. Daddi and colleagues ([118]), color–color diagrams based on the indices $z - K$ and $B - z$ can be used to detect and separate star-forming and passively evolving galaxies (at least) in the redshift range $1.4 < z < 2.5$ (Fig. 5.4). This method makes use of differences between the spectral energy distributions of star-forming galaxies and the SEDs of passively evolving galaxies near the rest-frame wavelength of 400 nm (see Fig. 6.3).

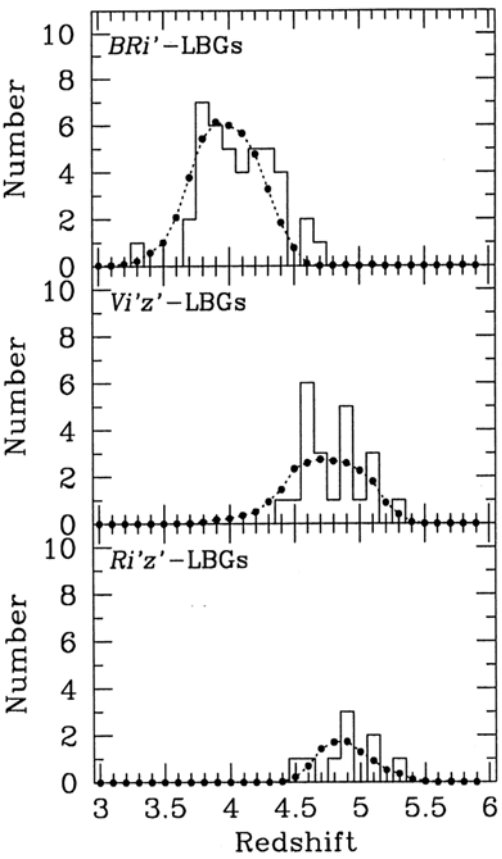

Fig. 5.3 Examples of predicted redshift distributions of Lyman-break galaxies observed in the Subaru Deep Field (see Sect. 5.2.3) with different filter-band combinations. The *dotted lines* show redshift distribution functions derived by means of Monte Carlo simulations. The histogram shows the distribution of redshifts of members of the corresponding samples for which spectra could be obtained. The distribution functions are approximately normalized to the histograms. (From [572])

In the literature, passively evolving high-redshift galaxies which have been identified by this method are referred to as "BzK galaxies." Similar selection schemes have been developed for identifying galaxies at lower redshifts ($1 < z < 3$). Depending on the specific criteria, such galaxies are referred to as FN, BX, and BM galaxies [5].

5.1.3 Identifications Using Emission Lines

The young hot stars of star-forming galaxies produce large amounts of UV continuum photons, which ionize the surrounding interstellar hydrogen. Recombination of the hydrogen produces Lyα photons. Due to the high absorption cross section

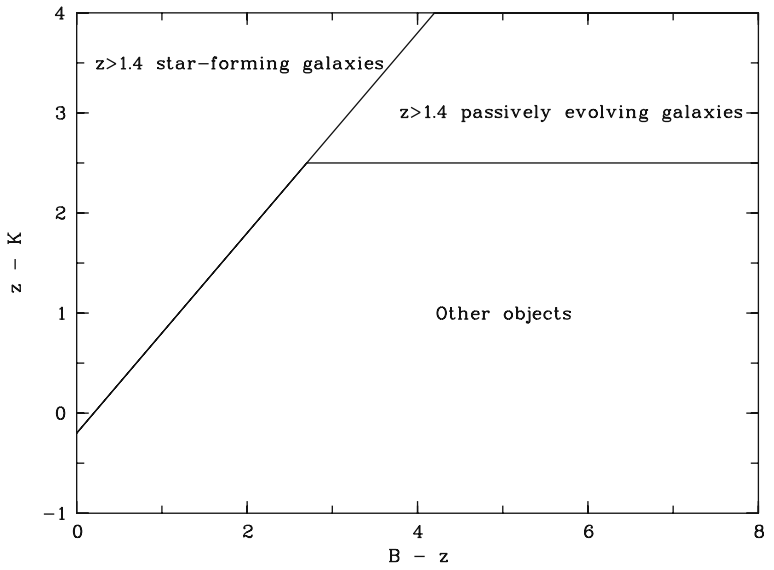

Fig. 5.4 Example of a BzK diagnostic diagram for identifying star-forming galaxies and passively evolving galaxies in the redshift range $1.4 < z < 2.5$ using B, z, and K photometric data. (For details see [118])

of hydrogen atoms for Lyα photons, the interstellar medium normally is opaque at this wavelength. Thus, the Lyα photons cannot directly escape. Instead, these photons experience multiple resonance scattering. If the galaxies are also opaque to Lyman-continuum photons, and if the UV photons are not destroyed by other processes (such as absorption by atoms or ions of other elements, or by dust absorption), the absorption and re-emission of the UV photons by the hydrogen atoms eventually converts all original Lyman-continuum photons into Lyα photons. If not internally absorbed, these photons eventually escape by diffusion in space and/or velocity space. Simple estimates show that in this way a significant fraction of the total luminosity of the hot stars can be converted into Lyα photons, and that in the absence of internal absorption Lyα emission equivalent widths of the order 200 Å can be expected for young star-forming galaxies. In this case, the flux density in the Lyα emission lines is expected to be much higher than in the adjacent continuum. As first pointed out by Partridge and Peebles in 1967 ([401]), this makes the detection of the line emission against the sky background much easier than the detection of the continuum emission of the high-redshift galaxies.

As noted in Sect. 1.3, strong Lyα emission indeed led to the first discovery of a galaxy with $z > 3$ by Djorgovski et al. [140] in 1985. But, as will be discussed in Sect. 7.1.3, the early searches for additional Lyα emitting galaxies were not successful, while many high-redshift starburst galaxies without Lyα emission were found. Obviously, in many star-forming, high-redshift galaxies the Lyα photons are reabsorbed inside the galaxies. However, this is not always the case. In 1998

a distinct population of distant starburst galaxies was found which shows intense Lyα emission (see Sect. 7.1.3). This subclass of the high-redshift starburst galaxies can be readily identified by their line emission, as originally suggested by Partridge and Peebles.

Some successful searches for high-redshift emission-line galaxies have been carried out using the objective grism technique and the Hubble Space Telescope [352, 568]. But most known high-redshift emission-line galaxies have been discovered from ground-based observations combining narrow-band and broad-band photometry. Normal galaxies, where the light is dominated by the stellar continua, tend to be visible with a similar brightness in many different filter bands longward of the (redshifted) Lyman continuum break (e.g., Fig. 5.5 c). As illustrated by Fig. 5.5

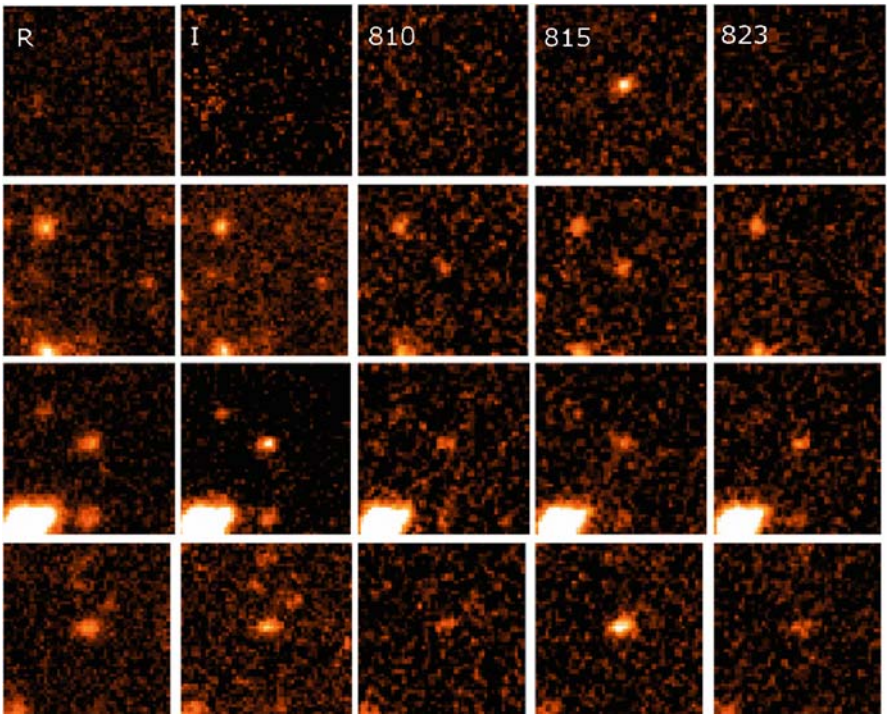

Fig. 5.5 The detection of Lyα emission line galaxies by means of narrow-band filter photometry. The figure shows images obtained with R and I broad-band and 810 nm, 815 nm, and 823 nm narrow-band filters. The *upper two panels* show Lyα galaxies with redshifts near $z = 5.7$. For the first galaxy (image **a**) significant flux is detected only in the 815 nm filter. The second galaxy (**b**) some flux also in the 810 nm band, since the Lyα line falls close to the boundary of two narrow-band filters. Galaxy (**c**) is a starburst galaxy at $z = 3.2$. At this redshift no strong emission line coincides with any of the narrow-band filters. Therefore, the flux observed in the five bands is of comparable strength. Galaxy (**d**) is a starburst galaxy at $z = 1.2$. In this case the redshifted 373 nm doublet of [O II] coincides with the 815 nm filter, producing a strong signal in this band. But the high flux in the R and I bands rules out Lyα as a possible identification of the observed line emission of this galaxy. (Adapted from [512])

a, b, and d, emission line objects are characterized by a conspicuous excess flux in one or (if the emission line coincides with the overlap region of two filters) in two adjacent filter bands. Since the continuum contribution decreases proportional to the filter band width, while the emission line contribution of a narrow line remains constant, the effect is particularly strong if a narrow emission line is observed in a filter band of a comparable width. In this way filter photometry allows for an efficient and sensitive detection of galaxies which have emission lines in their spectra.

More problematic is the unambiguous identification of the observed emission lines. Since, as pointed out earlier, all sky surveys are dominated by foreground galaxies, the great majority of the emission features detected on objective grism frames and in photometric emission-line searches are [OII], [OIII], Hα, and other common interstellar emission lines produced by foreground galaxies with moderate redshifts. These foreground objects have to be eliminated by means of photometric redshifts. For this purpose the narrow-band photometry has to be supplemented by deep broadband images of the same field. Since the foreground objects normally are detected in many different filters, their photometric redshift generally can be derived with sufficient accuracy. However, galaxies with highly redshifted Lyα lines often are "dropouts" at all optical broad-band filters, and photometric redshifts based on NIR data alone are not always reliable. Thus, the true nature of apparent high-redshift Lyα emitters normally has to be confirmed by means of spectroscopic observations. With spectroscopic data, misidentifications can be ruled out by means of the observed line profiles (see Sect. 7.1.3) and by the absence of other emission lines, which are always present in the spectra of lower-redshift emission line objects.

The best signal-to-noise ratio in emission-object searches are to be expected in filter bands corresponding to wavelength regions of low night-sky background. Since most of the night-sky background in the red and near infrared is due to airglow produced by OH radicals in the high atmosphere, optimal filter bands correspond to gaps or "windows" in the OH line spectrum. The presence and location of such OH windows is clearly visible in the night-sky spectrum reproduced in Fig. 4.9. The central wavelengths λ_c and the corresponding Lyα redshifts for some OH-windows are listed in Table 5.1. Typical widths of the OH windows are in the range of 30–70 nm. Emission-line searches use either narrow-band filters covering a complete OH window, or several narrower, adjacent filter bands inside an OH window. By using several filters residual OH emission inside the window can be avoided, and significantly higher sensitivities can be reached. An example of such a filter set is given in Fig. 5.6. In this set the filter centered at 8240 Å completely avoids significant air-glow lines. The 8160 Å filter includes some lines, but only in the wings of the transmission function. Only for the 8100 Å filter a significant contamination by residual OH could not be avoided. As expected, observations in the 8240 Å filter reach the deepest magnitudes, while the 8100 Å observations have

Table 5.1 Atmospheric OH-emission windows in the red and NIR spectral range suitable for detecting distant Lyα emitters at redshifts $4 < z < 10$

λ_c (nm):	670	815	910	1062	1080	1112	1182	1316	1336
z(Lyα):	4.5	5.7	6.5	7.7	7.9	8.1	8.7	9.8	10.0

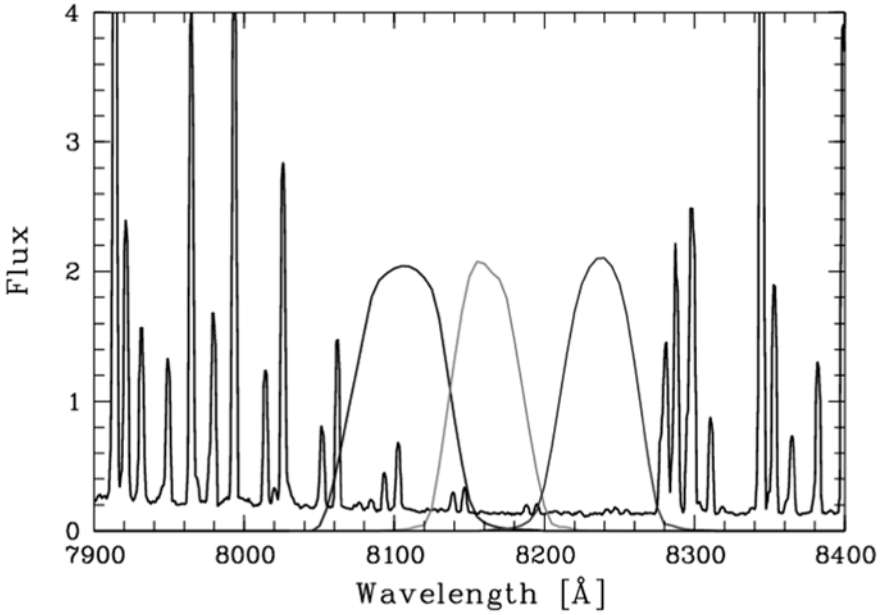

Fig. 5.6 Night-sky OH emission spectrum (*thick lines*) and the transmission curves of a typical narrow-band filter set (*thin lines*) designed for searches of Lyα emission objects near $z \approx 5.7$. In order to facilitate the detection of faint emission sources, the filters are designed to include a minimum of night-sky flux. (From Tapken et al. [512])

the lowest sensitivity. But in all three filters lower line fluxes are reached than is possible with a single filter covering the whole OH window.

An additional potential advantage of using several filter bands inside an OH window is that the flux ratios between the adjacent filters (instead of the flux in a broad-band filter) can be used for determining the underlying continuum flux (or its upper limit). Since fewer photons are recorded, the narrow-band filters are less well suited for accurately measuring the continuum flux. But adjacent narrow-band filters give more reliable information on the origin of the excess emission in a narrow filter band and on the emission-line character of this excess.

Lyα emission is also present in the spectra of AGN. Therefore, during searches of emission line objects, in addition to emission-line galaxies, AGN are detected. However, the space density of AGN is much lower than that of galaxies. Therefore, the contamination of emission-line galaxy samples by AGN typically is not more than a few percent. Moreover, AGN usually can be effectively separated from the galaxies by means of their higher X-ray and radio emission.

5.2 Search Strategies

Characterizing the typical properties of high-redshift galaxies require statistically meaningful samples of such objects. In order to find and to identify such sam-

ples, the identification techniques described in Sect. 5.1 have to be applied to representative cosmic volumes. The volume sampled by a given program depends on the observed area of the sky and on the depth of the observations. Obviously a combination of a large sky area and a high depth would be ideal. In practice it is difficult to achieve a good sky coverage and a high depth simultaneously. Reaching deep into space means observing very faint objects. This requires long integration times at very large telescopes. Because of the high demand for time at large telescopes and since such telescopes are used for many different scientific programs, very deep imaging or spectroscopy on large telescopes usually is limited to few and relatively small fields. In order to observe larger areas of the sky, smaller telescopes have to be used, which can be dedicated for such surveys and which can be made available for this single purpose. But the smaller telescopes cannot reach as deep as the largest available instruments. On the other hand, photometric observations with medium size ($\sim$ 4-m) telescopes are well suited to identify high-redshift objects in the magnitude range which is accessible for spectroscopy with 8-m or 10-m telescopes.

"Shallow" large area surveys and very deep surveys in small fields often complement each other. Both can sample similarly large volumes. However, these volumes are of different nature. The depth which can be reached in a survey obviously depends on the absolute luminosity of the observed objects. Therefore, even moderately sensitive large-area surveys reach a high depth (and consequently a large volume) for intrinsically very luminous objects. Since very luminous objects are rare, they are, in fact, found most easily in large-area surveys. Therefore, in addition to providing complete samples for galaxies in the local universe and at low redshifts, large area surveys allow us to find rare very luminous objects (such as high-luminosity QSOs) at high redshifts. On the other hand, deep surveys with large telescopes (usually referred to as "deep fields") allow us to sample the more common moderately luminous high-redshift galaxies. Since the space density of these more common objects is high, deep surveys covering modest volumes can give statistically significant data.

The past and current astronomical surveys cover a range of field sizes and depths. Somewhat arbitrarily, in the following they will be sorted into three categories. First we will discuss large-area surveys which cover the entire sky, or at least a major fraction of the whole sky. For the reasons outlined above these large-area surveys are relatively shallow. A second section will discuss medium deep surveys with fields of the order 1–50 square degrees. The third section will be devoted to "deep fields," which, naturally, are most important for finding the high-redshift galaxies. Figure 5.7 shows the observed area and the limiting I-band magnitude in graphic form for some examples of recent photometric surveys.

The final section of this chapter will discuss the "Gravitational Telescope" technique, which makes use of the flux amplification of distant galaxies by the gravitational field of foreground mass concentrations. This allows us to observe distant galaxies which otherwise would be out of reach of our present telescopes, and which cannot be reached with normal deep field techniques.

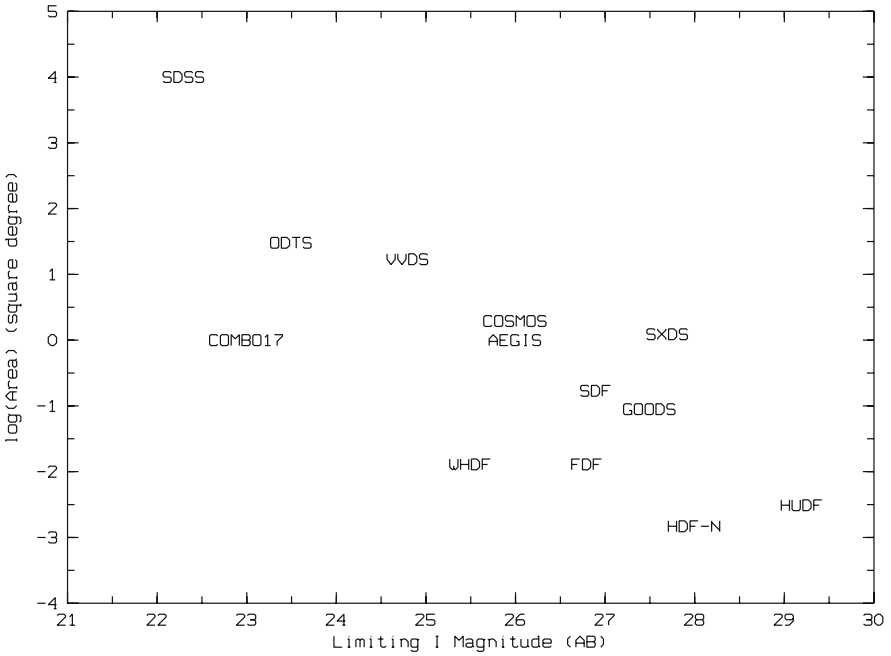

Fig. 5.7 Observed areas and limiting (AB) I-band magnitudes of some examples of recent large galaxy surveys. The VVDS data point represents the CFHT deep imaging survey [306], which formed the basis of the current VVDS spectroscopic survey. For programs with several different subprojects, the data for the largest field with I-band photometry are given

Because of the very large number of past, current, and future surveys, it is not possible to list or to describe all existing or planned surveys. Moreover, while important and valuable at the time of their execution, due to the progress in astronomical technology, surveys tend to become superseded by new, larger, deeper, or more comprehensive surveys on relatively short timescales. Therefore, only general properties and some typical examples will be discussed here. Informations on many additional (current and future) astronomical survey programs can be obtained from the web site of the "Working Group for Sky Surveys" of the International Astronomical Union (http://www.skysurveys.org) and the web sites of the large international observatories.

5.2.1 Large-Area Surveys

Sky surveys have been carried out since the beginning of astronomy. The first surveys covering the whole sky were initiated during the 19th century by combining observations from observatories in the northern and the southern hemisphere.

Probably the best-known large-area optical sky survey is the "Palomar Observatory Sky Survey" (**POSS**), which was originally carried out between 1948 and

1958 using photographic imaging in two filter bands with a dedicated (Schmidt-type) telescope on Mt. Palomar in Southern California. Between about 1989 and 1999 the original POSS observations were repeated as POSS II with new and better photographic material in three colors (for details on the POSS surveys see [430]). Originally, the POSS covered only the part of the sky visible from Mt. Palomar. Starting in 1973 the POSS was supplemented by Schmidt-telescope observations with similar properties carried out at the La Silla site of the European Southern Observatory (ESO) in Chile and the SRC Schmidt telescope at the Siding Spring Observatory in Australia. Taken together this POSS–ESO–SRC survey covers the whole sky. The latest versions of these surveys have been digitized and are publicly available through the web pages of the STScI and the major astronomical data archives (such as the CDS, see Sect. 4.4). Moreover, the objects recorded by the POSS and their positions are listed in various digital catalogs (e.g., [381]).

The Palomar-ESO-SRC Schmidt surveys played a decisive role in the development of extragalactic astronomy. In the POSS the sponge-like large-scale structure of the galaxy distribution was observed for the first time, significant samples of galaxy clusters could be derived, and many radio galaxies and QSOs were identified using POSS plates. In spite of its (as seen from today) modest limiting magnitudes ($B \approx 22.5,\ R \approx 20.8$), the Schmidt sky surveys reached well into the high-redshift universe by recording the images of many $z > 2$ QSOs. For the reasons discussed in Sect. 4.2, these photographic surveys do not reach the faint radiation fluxes of high-redshift galaxies. But, because of their completeness, they still form an important reference source for optical astronomy.

Among the most important modern large-area surveys at visual wavelengths is the first part of the "Sloan Digital Sky Survey" (**SDSS**). SDSS-1 was carried out between 1998 and 2005 using a dedicated 2.5-m telescope at the Apache Point Observatory in New Mexico. It covered about 10^4 square degree (or about one quarter of the whole sky) and provided photometry in 5 optical filterbands (down to about a visual magnitude of $m = 23$) for about 2×10^8 objects. Moreover, using a fiber-type multi-object spectrograph, spectra of about 675 000 galaxies, 90 000 QSOs, and 185 000 stars brighter than $m_r \approx 18.2$ were observed. (For details see [569], or http://www.sdss.org.)

Apart from providing fundamental statistical data for the local universe, the SDSS increased the number of known QSOs by about a factor of three. Moreover, the SDSS resulted in the first discovery of QSOs with redshifts $z > 6$. However, like the other large-area surveys mentioned above, the SDSS is to shallow to include normal high-redshift galaxies with $z > 2$.

All-sky surveys also exist for the infrared wavelength range and for the X-ray sky. Among these full-sky maps are the NIR "Two Micron All Sky Survey" (**2MASS**, [479], supplemented by the southern NIR survey **DENIS** [158]), and the X-ray "ROSAT All Sky Survey" (**RASS**, [528]). An important first survey of the FIR sky was carried out with the Infrared Astronomical Satellite (**IRAS**) already in 1983 [489]. IRAS carried out photometry in four bands (12, 25, 60, and 100 μm) and covered about 96% of the sky.

A new large-area NIR survey is being carried as part of the "UKIRT Infrared Deep Sky Survey" (**UKIDSS**, [303]). UKIDSS covers about 7000 square degrees of the northern sky in five NIR filter bands (Z, Y, J, H, K) with a limiting magnitude of $m_K \approx 18$ (compared to $m_K \approx 14.3$ for 2MASS).

At radio wavelengths the "NRAO VLA Sky Survey" (**NVSS**, [108]) covers the radio sky north of -40° declination. NVSS is a 1.4 GHz continuum survey with a completeness limit of about 2.5 mJy and a position accuracy of a few arcsec. A slightly greater depth (detection limit ≈ 1 mJy) and a somewhat higher positional accuracy is reached by the "Faint Images of the Radio Sky at Twenty-cm" survey (**FIRST**, [34]), which covers about 10^4 square degrees near the galactic poles. All other large-area radio surveys have significantly lower resolution and lower positional accuracy, making it difficult to identify the observed radio sources with distant galaxies.

A first large-area survey for the extragalactic far and extreme ultraviolet (FUV and EUV) sky has been initiated with the "Galaxy Evolution Explorer" experiment (**GALEX**), [338]) which was launched in 2003. GALEX aims at FUV/EUV imaging and photometry of all galaxies accessible to its relatively small (50 cm aperture) space-based telescope. Technical data and information on already released data and on the project status can be found in the project web page (http://www.galex.caltech.edu). Providing for the first time extensive FUV and EUV data for extragalactic objects, this experiment had an important impact on our understanding of the local and moderately distant universe. But its operating wavelength obviously is not suited to study high-redshift objects.

5.2.2 Medium Deep Surveys

The published medium deep surveys typically reach galaxies up to redshifts $1 < z < 2$. But, due to their relatively large volume, they often include some more distant objects with particularly high luminosities. Among the notable examples for medium deep surveys is again the NIR UKIRT Infrared Deep Sky Survey (UKIDSS, see Sect. 5.2.1) which includes a medium deep component, covering about 35 square degrees with a limiting K-magnitude $m_K \approx 21$. In the thermal infrared the "Spitzer Wide-area InfraRed Extragalactic" survey (**SWIRE**, [320], http://swire.ipac.caltech.edu) has mapped about 50 square degrees of the sky using the IRAC and MIPS cameras of the NASA infrared satellite Spitzer. SWIRE includes L^* galaxies with redshifts up to about $z = 2$ and more luminous galaxies at even higher redshifts.

An important optical medium-deep survey is the "Classifying Objects by Medium-Band Observations in 17 Filters" (**COMBO17**) imaging program ([563], see also Fig. 6.1), which covered ≈ 1 square degree in three fields of the southern sky. The survey resulted in redshifts, SEDs and other parameters for 25 000 galaxies with $m_R \leq 24$ in the redshift range $0 < z < 1.55$. It was primarily designed to study galaxy evolution in the range $0 < z < 1.2$. But the COMBO17 images almost certainly also contain (not yet identified) higher redshift galaxies.

COMBO17 observations also cover the Extended Chandra Deep Field (ECDF, see Sect. 5.2.3), where more than 63 501 objects were observed [564]. An important extension of this program is the ≈ 900 square arcmin "Galaxy Evolution From Morphology and SEDs" (**GEMS**, [83, 438]) survey, which combines the COMBO17 observations of the ECDF region with morphological information for about 10^4 galaxies derived from HST images of the same field.

Another recently completed medium deep photometric program is the "Oxford-Dartmouth Thirty Degree Survey" (**ODTS**, [325], which covers about 30 square degrees in 5 or 6 optical bands and, in part, the NIR K-band to a V magnitude of about $m_V = 25$).

An ongoing medium deep project is the "Multiwavelength Survey by Yale-Chile" (**MUSYC**) (actually an international program involving institutes from the US, Chile, and Europe), which covers four southern or equatorial fields with a total surface of 1.2 square degrees (see, e.g., [197] and [522]). MUSYC combines optical and IR data with X-ray, UV, and radio imaging. The imaging program is being supplemented by spectroscopic observations being carried out at the ESO VLT and the Baade Telescope of the Las Campanas Observatory in Chile.

A cooperation of several Spanish institutes has recently started the **ALHAMBRA** project [361]. This survey includes photometry in 20 medium band filters covering the spectral range of 350–970 nm, plus NIR photometry in J, H, and K of a 4 square degree area. The flux limit is about 25 magnitudes in the optical bands and about 20 magnitudes at the K-band. The survey is expected to yield photometric redshifts of a sample of about 10^5 galaxies with a median redshift of $z \approx 0.7$.

Among the recent or ongoing spectroscopic medium-deep surveys are the Deep Extragalactic Evolutionary Probe **DEEP2**, the **zCOSMOS-bright**, and the VIMOS VLT Deep Survey (**VVDS**) projects. DEEP2 covers about 3.5 square degrees and provides 40 000 medium resolution spectra and spectroscopic redshifts for $0.7 < z < 1.4$ galaxies ([121]). The zCOSMOS-bright program (e.g., [314]) covers about 2 square degrees (which are also imaged with the HST and with groundbased telescopes) and aims at obtaining low-resolution spectra and redshifts of about 4×10^4 galaxies with $I_{AB} < 22.5$ in the redshift range $0.2 < z < 1.2$. The VVDS will eventually yield low-resolution spectra and redshifts of about 10^5 galaxies ([307]). The targets for the VVDS were selected from a UBRI photometric survey covering about 17 square degree, which was carried out at the Canada–France–Hawaii Telescope (CFHT) at Mauna Kea, Hawaii. In the I-band the limiting (AB) magnitude of this photometric survey is about 25.3 [306].

Currently, many different medium-deep surveys are being executed, prepared, or planned. Their results, which will become available within the coming decade, promise important statistical data (mainly) for the intermediate redshift galaxies.

5.2.3 Deep Fields

The up-to-now deepest look into the cosmos at optical wavelengths has been the "Hubble Ultra Deep Field" (**HUDF**, [35], Fig. 5.8) observed in 2003 with the

Fig. 5.8 The Hubble Ultra Deep Field (HUDF). (Credit: NASA, ESA, and S. Beckwith (STScI) and the HUDF Team)

Advanced Camera for Surveys (ACS) of the Hubble Space Telescope. The HUDF covers a quadratic field of about 11 square arcmin, centered on RA = $3^h\ 32^m\ 39.0^s$, Decl. = -27^0 47' 29.1" (J2000). Four optical filter bands were observed with a total integration time of about 11.2 days. In all four bands the limiting (AB) magnitude was at least $m = 29$ (corresponding to flux levels near 10 nJy). Most of the about 10^4 objects detected in the HUDF are distant galaxies.

In parallel to the HUDF, the HST observed two smaller fields of 0.7 square arcmin with the Near-Infrared Camera and Multi-Object Spectrometer (NICMOS) of the HST. The resulting NICMOS frames form the deepest NIR images available so far. As an extension of the original UDF program the two NICMOS parallel fields have later been imaged in three optical filter bands with the ACS camera ("UDF05," [382]). The corresponding ACS frames are about one magnitude less deep than the HUDF observations, but are also among the deepest looks into the cosmos. Due to the failure of the ACS in 2006, the UDF05 program could be only partially completed.

The NICMOS NIR camera was also used to image an about 6 square arcmin subfield of the HUDF at 1.1 and 1.6 μm with a limiting (AB) magnitude of $\approx$ 27.7 ([520]). Deep optical frames obtained with the ACS in parallel to these NICMOS observations are referred to as the "HUDF Parallel Fields" (or **HUDF-Ps**). The HUDF-Ps images are about 0.7–1.0 magnitudes less deep than the HUDF(e.g.,[64]).

Two earlier very deep optical images obtained with the HST are known as the "Hubble Deep Fields North" and "South" (**HDF-N, HDF-S**, [558], [557], [170]).

Each of the HDFs covers an area of 5.3 square arcmin, and (like the HUDF) was imaged in four optical filter bands. The HDF-N is centered on RA = $12^h\ 36^m\ 49.4^s$, Decl. = $+62^0$ 12' 58" (J2000), the HDF-S on RA = $22^h\ 32^m\ 56.2^s$, Decl. = -60^0 33' 02.7" (J2000). Both fields had limiting magnitude of about $m = 28$ in the visual spectral range, and each field recorded about 3000 objects. The HUDF and the HDFs resulted in a large number of follow-up studies at many different wavelengths and by many different groups.

Early ground-based deep field observations were carried out using 4-m class telescopes. Some of these projects reached rather faint limiting magnitudes. A notable example is the William Herschel Deep Field **WHDF**, which was observed using the William Herschel telescope of the Roque de Muchachos Observatory on La Palma, Spain. The WHDF covered a 49 square arcmin area and reached galaxies down to a blue magnitude of about 28.5 [350].

Of greater impact have been deep-field observations carried out with the new generation of >8-m telescopes, which (due to their active optics systems and shorter exposure times) could produce images of better quality and better angular resolution.

One of the first groundbased surveys carried out with a ≥8-m telescope has been the "FORS Deep Field" (**FDF**, [16, 228]), which was observed (mainly) in 1999 using the FORS instruments of the ESO VLT for the optical bands and the 3.5-m ESO NTT for the NIR bands. The FDF, centered on RA = $1^h\ 06^m\ 03.6^s$, Decl. = -25^0 45' 46" (J2000), covers about 46 square arcmin in nine filter bands (U, B, g, R, I, 834 nm, z, J, and K). At the B-band the FDF reached a limiting magnitude only 0.3 mag lower than that of the HDF-N, with an integration time <20% of the time invested for the HDF-N B-band observations.

The FDF demonstrated that (at least) in the blue and visual spectral range groundbased imaging can reach as deep as space observations. In the red and NIR ranges the atmospheric OH emission makes it more difficult and more time consuming to match the depth of space-based deep fields from the ground. Moreover, at present, the angular resolution of wide-field ground-based observations is still significantly inferior to the images obtainable from space. On the other hand, the shorter exposure times with the large ground-based telescopes makes it possible to observe more filter bands and to derive more spectral information on the observed targets. Present groundbased telescopes allow us to observe larger fields, and consequently larger volumes of space in a single frame. The galaxy sample observed in the FDF, e.g., contains about the same number of objects as the HUDF (and about three times more objects than the HDFs).

A deep field project aiming at the identification of high-redshift galaxies, carried out with the Keck I telescope, has become known as the **Keck Deep Fields**. This program includes imaging in four filter bands to a limiting magnitude $R \approx 27$. (For details see [442].)

As noted already, an advantage for groundbased deep-field programs is the availability of larger cameras and of multi-object spectrographs covering larger areas. A prominent example of a deep-field project making use of this advantage is the "Subaru Deep Field" (**SDF**, [259]) which was observed with the 34' × 27'

Suprime-Cam camera of the Japanese national telescope Subaru. The SDF, centered on RA = $13^h\ 24^m\ 38.9^s$, Decl. = $+27^0$ 29' 25.9" (J2000), covers 725 square arcmin in seven filter bands. Its depth is again comparable to that of the original HDFs. Another deep field based on Subaru observations is the "Subaru/XMM-Newton Deep Survey" (**SXDS**, [188, 465]). The SXDS covers five bands (B, V, R, i', z') and reaches the AB magnitudes B =28.4, V = 27.8 R = 27.7, i' = 27.7, and z' = 26.6. It covers 1.2 square degrees. The optical imaging has been supplemented by deep observations at radio and X-ray wavelengths. The SXDS is centered on RA 2^h $18^m\ 00^s$, DEC -5^0 00' 00".

Several current deep-field projects combine deep observations in different wavelength ranges, using coordinated observations with different telescopes and satellites. A particularly interesting example is the "All-Wavelength Extended Groth Strip International Survey" (**AEGIS**, [122]) which covers a 2 degree long narrow strip of the northern sky near RA = $14^h\ 17^m$, Decl. = $+52^0$ 30'. Compared to a quadratic (or nearly quadratic) field of the same surface, a strip has the advantage of covering a larger range of linear scales. Therefore, statistical conclusions from a strip-type deep field are less affected by the "cosmic variance" due to the large-scale structure of the universe. AEGIS combines deep observations at X-ray, FUV, optical, FIR, MIR, and radio wavebands. The optical observations include HST imaging, ground-based observations with the Canada–France–Hawaii Telescope, and optical spectroscopy with the Keck telescope. The areas covered in the different bands range between 0.2 and 1 square degree.

A particularly deep multi-wavelength survey is the Great Observatories Origins Deep Survey (**GOODS**, [132, 204]), which combines ground-based optical, submillimeter, and radio observations (including spectra obtained at the Keck, VLT, and Gemini telescopes) with observations obtained with the HST and with the Spitzer IR satellite and the Chandra X-ray space observatory. GOODS covers two fields of 10' × 16' which are centered on the HDF-N and on the X-ray "Chandra Deep Field South" (CDFS, see below).

Another project combining large-area and high depth is the "Cosmic Evolution Survey" (**COSMOS**, [464]), which promises important new information on the evolution and space distribution of galaxies. COSMOS covers a $1.4^0 \times 1.4^0$ field at RA = $10^h\ 00^m\ 28.6^s$, Decl. = $+02^0$ 12' 21.0" (J2000), part of which will be investigated to a depth comparable to the deep fields mentioned above. For this purpose deep optical imaging is being obtained with the ESO VLT, the Subaru telescope, and the CFHT. For FUV imaging the GALEX satellite is used, and deep X-ray imaging is provided by the XMM-Newton and the Chandra X-ray satellites. Infrared images are being contributed by the Spitzer satellite, the UKIRT facility, the CTIO 4-m telescope, and the University of Hawaii observatory. Submillimeter and radio data are collected at IRAM, at the JCMT, the VLA, and the Sunyaev-Zeldovich Array (SZA). Spectra for >15 000 of the $\approx 3 \times 10^6$ observed galaxies are obtained, (mainly) with the ESO VLT and the Magellan telescopes of the Las Campanas observatory in Chile. An overview of the program and some first results have been presented in a series of articles in the special issue Vol. 172 (Sept. 2007) of the *Astrophysical Journal Supplement Series*.

Very deep NIR photometry (with limiting (AB) magnitudes of 26.8, 26.2, and 26.2 at, respectively, the J, H, and K bands) has been obtained as part of the **FIRES** program [294]. This survey covers a 2.5' × 2.5' field centered on the HDF-S. The observations have been carried out with the ISAAC camera of the ESO VLT. To reach an optimal sensitivity all IR images have been obtained under better than average seeing conditions.

A deep-field component (in addition to the large-area and medium-deep surveys mentioned above) is also included in the UKIDSS project. The **UKIDSS** "Ultra Deep Survey" covers an area of 0.77 square degree and will reach $m_J = 24.8$ and $m_K = 22.8$.

A valuable spectroscopic survey based on deep NIR data is the **K 20** survey ([353]) which yielded about 500 spectra of NIR selected galaxies with redshifts $z \leq 2.7$.

Deep-field projects have also been carried out for the X-ray sky and at submillimeter wavelengths. The more recent X-ray deep fields have been reviewed by Brandt and Hasinger [70], who list 22 deep extragalactic X-ray surveys carried out with either NASA's Chandra satellite or the XMM-Newton instrument of ESA. The most sensitive of these surveys are the "Chandra Deep Fields" **CDFN** [13]) and **CDFS**, [202]), each covering about 0.1 square degree. The CDFN is centered approximately on the HDF-N and has for 0.5–2 keV photons a detection limit of about 0.007 nJy. The CDFS (RA = $03^h\ 32^m\ 25^s$, Decl. = -27^0 48' 50" (J2000)) has a flux limit of about 0.014 nJy. The CDF-S is surrounded by the "Extended CDF-S" field which has been imaged with a somewhat lower sensitivity.

Among the deep fields observed at submm wavelengths is the "Canada-UK Deep Submillimeter Survey" (**CUDSS**, e.g., [553]), which covers two 6'×8' fields. This survey was carried out using the SCUBA bolometer array at the JCMT submm telescope in Hawaii. The wavelength was 0.850 mm and the flux limit about 3 mJy. The same telescope has been used for the somewhat larger, but less sensitive (flux limit ≈ 8 mJy) "SCUBA Half-Degree Extragalactic Survey" (**SHADES**, [364]).

5.2.4 Searches Using "Gravitational Telescopes"

Recent deep surveys have been highly successful in detecting samples of distant galaxies and QSOs. However, at very high redshifts such programs usually reach only the brightest objects of their kind. These exceptionally bright objects may not be representative of their respective populations. On the other hand, with present telescopes, it is difficult to extend the luminosity limits of deep surveys significantly. According to Eq. (4.10), for faint objects, where the photon statistics of the sky background dominates the noise, the exposure time required to reach a given S/N increases with decreasing object flux $\sim F^{-2}$. Some of the recent deep fields are based on observing programs with total exposure times of days or weeks. Extending such programs to objects which are fainter by a factor of ten requires (at least) 100 times more observing time. Obviously, such programs are unrealistic.

On the other hand, it is possible to observe such intrinsically faint or very distant sources with acceptable observing time, if their flux is amplified by the gravitational lensing effect of foreground masses (Sect. 2.3). For distant galaxies lensed by foreground galaxies or galaxy clusters, amplification factors >10 are not uncommon (e.g., [184, 346, 449, 481]). In the literature the gravitational flux amplification of faint distant objects often is referred to as the "gravitational telescope technique" [407]. Obviously, this technique is an efficient and economic way to reach deeper into space with the present generation of instruments.

Since the sky area with high flux amplification factors are larger around galaxy clusters then around galaxies, clusters are better suited as gravitational telescopes. Numerical estimates show that if high-redshift galaxies are observed behind a galaxy cluster of high-lensing strength (such as the cluster corresponding to the ROSAT X-ray source RX J147-1145) a flux amplification by factors between 2 and 10 occurs in an area which is larger than the HUDF.

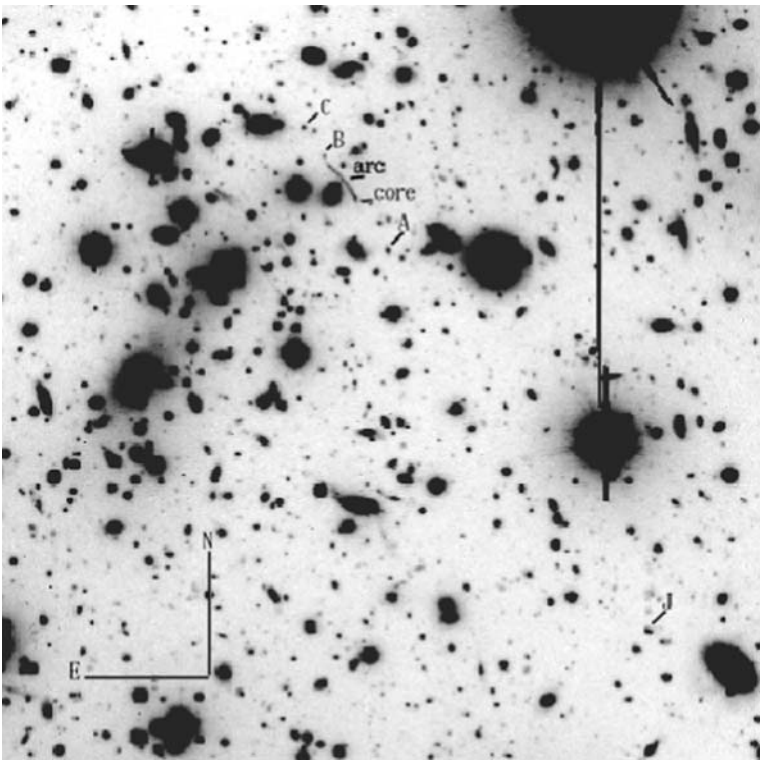

Fig. 5.9 Examples of gravitationally amplified distant galaxies. The VLT image shows part of the X-ray bright $z = 0.296$ galaxy cluster 1E0657-56. The narrow gravitational arc in the upper half of the image and the galaxies labeled C, B, A, and J (all oriented along a *straight line*) are gravitationally lensed distant background galaxies with redshifts $2.19 < z < 3.24$ which are gravitationally amplified by factors between 2.5 and > 20. (From Mehlert et al. [346])

An example for the efficiency of this technique is an investigation of the X-ray bright galaxy cluster 1E0657-56 (Fig. 5.9) by D. Mehlert et al. [346] based on a single long-slit spectrum. This spectrum was obtained as a short test exposure during the commissioning of one of the first instruments at the ESO VLT in 1998. As shown in Fig. 5.9 the region of this cluster contains several prominent arcs, which are distorted and amplified images of galaxies behind the $z = 0.296$ cluster. The long-slit spectroscopic test exposure was made to get a spectrum of the narrow "arc" in Fig. 5.9. As expected the arc was found to be the lensed image of a distant ($z = 3.24$) galaxy. Four other objects (labeled A, B, C, and J in Fig. 5.9) which were included in the ≈ 7 arcmin long slit turned out to be high-redshift ($z > 2$) galaxies as well. A flux magnification map of the cluster derived from the various

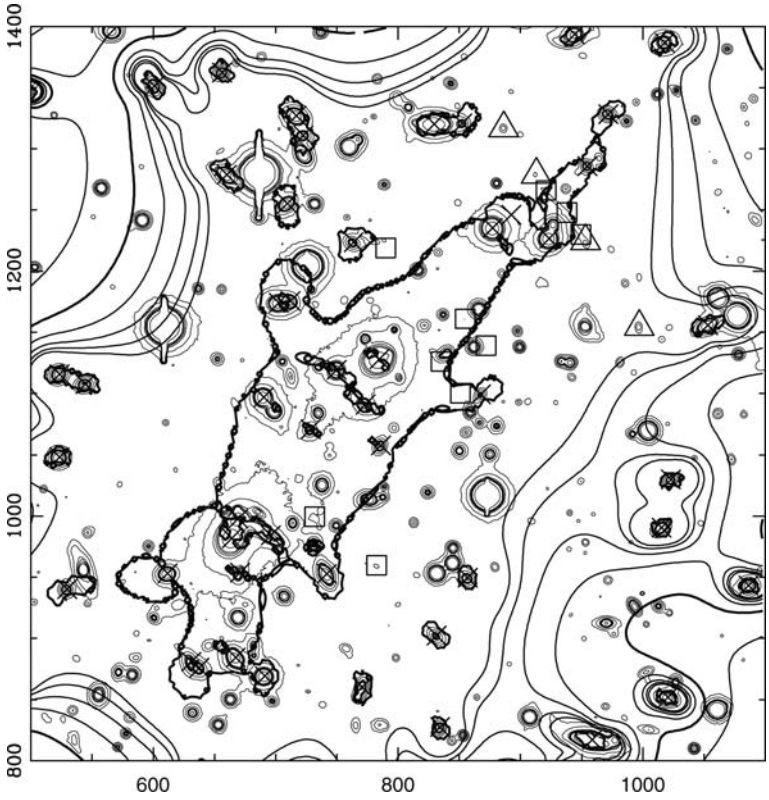

Fig. 5.10 Map of the amplification factors for background galaxies with a redshift of $z = 3.24$ behind the galaxy cluster 1E0657-56. The outer *thick solid line* corresponds to a computed magnification factor of 6. The central *thick solid contour* corresponds to the "critical line" where the magnification formally becomes infinite. The *dotted lines* between the *two solid lines* are the contours with magnification factors of 7, 8, 9, and 10, respectively. The outer *thin*, *dotted lines* are the contours with amplification factors of 5 and 4. The *open triangles* indicate the positions of the lensed galaxy images C, B, arc, and A (shown in Fig. 5.9). The galaxy J is outside the map. (From Mehlert et al. [346])

arcs and lensed images is reproduced in Fig 5.10. It shows that all five observed lensed objects are significantly amplified. The amplification factors range between 2.5 (for Object J) and about 21 (for the arc).

Not all galaxy clusters are sufficiently concentrated to make good gravitational telescopes. However, many clusters suitable as gravitational telescopes are known. So far only a few have been used to observe deeper into space. Although this technique has already been profitably used to discover and to study a few highly magnified distant galaxies (e.g., [68]), the great potential of the gravitational amplification has not been fully exploited so far. At least in part this seems to be due to the fact that corresponding proposals are not much appreciated by the program committees of the existing large telescope facilities.

Chapter 6
Deriving Physical Parameters

Determining the physical properties of distant galaxies is of fundamental importance for understanding the high-redshift universe. Firstly, measuring the galaxy properties as a function of redshift (or cosmic age) provides important information on the poorly understood nonlinear processes governing the evolution of galaxies. Secondly, mapping the space distribution of matter at different redshifts for a comparison with the predictions of the cosmic structure formation theory requires methods to derive the mass of galaxies from the observational data. This cannot be achieved without understanding the basic properties of these objects. Finally, without some knowledge of the galaxy properties and their statistics, the detection probabilities of distant galaxies and the completeness of surveys cannot be estimated.

If the redshifts are known, some galaxy properties (such as the galaxy sizes and morphologies) can be derived (more or less) directly from images of sufficient angular resolution. Most other parameters have to be deduced from spectra or spectral energy distributions. Using SEDs derived from photometry makes it possible to study larger samples and to reach deeper, but spectroscopic observations give more detailed, more accurate, and more reliable results.

6.1 Geometrical and Kinematical Properties

Among the most basic parameters of galaxies is their size. For nearby galaxies the surface brightness can be mapped and the shape and the dimensions can be accurately described. In the case of distant galaxies, the $\sim (1+z)^{-4}$ dimming of the surface brightness (Sect. 3.1.5) makes it difficult to observe the outer regions of the galaxies against the sky background. Moreover, even on space-based images the very distant galaxies often are not well resolved. Therefore, for comparing galaxy sizes, a parameterization is needed which can be used at all distances. The most common measure of galaxy sizes is the radius around the light center inside of which 50% of the total flux is emitted. If the surface brightness as a function of radius is known or can be measured, this "half-light" or "effective" radius r_e (Sect. 2.2.2) can be derived by means of a model fit to the observable part of the light distribution. The total flux can then be derived by extrapolating the observed part of the surface

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_6,

brightness outward to the distance where the missing flux becomes negligible. This is possible, since the surface brightness of galaxies decreases steeply and follows relatively simple and well-known laws (see Sect. 2.2.1). If the distance of a galaxy is known, r_e derived in arcsec can be directly converted into an effective radius in kpc.

Because of the different colors of the central bulges and of the disks, the effective radii of disk galaxies normally depend on the filter band of the observations. Therefore, for comparisons of radii of galaxies at different redshifts, images corresponding to the same rest-wavelength should be used. The systematic errors introduced by observations in different rest-frame bands depend on the galaxy type and the redshift. Studies of UV-bright high-redshift galaxies are facilitated by a rather small dependence of the size on the (UV to visual) observing wavelength (see, e.g., [203] and [400]). The absence of a significant color dependence of the radius of these galaxies probably results from the relatively uniform (young) stellar population of these objects.

All information on the kinematics of galaxies results from spectroscopic Doppler effect measurements. As pointed out before, in principle, Doppler effect measurements have the great advantage of being independent of the distance. However, in practice the small size of the high-redshift galaxies makes kinematic studies increasingly difficult with increasing redshift. Because of the required high light collecting power, normally ground-based telescopes are used for spectroscopy. Spectra obtained with these instruments often are blurred by seeing effects. Moreover, often spectrograph slit widths comparable to the galaxy sizes have to be used. Therefore, spatial information on the kinematics of high-redshift galaxies is either completely lost or has to be reconstructed from the unsharp data. This is usually done by modeling the blurring effects and by reconstructing the original spectra using suitable models based on high-resolution imaging data of the same field. (For details and examples see, e.g. [54].)

6.2 Redshifts and Distances

6.2.1 Measuring Redshifts

According to Eq. (1.2) the redshift z is defined by the relation $z + 1 = \lambda_{obs}/\lambda_{em}$, where λ_{obs} is the observed and λ_{em} is the emitted wavelength. z is traditionally measured by comparing the wavelengths of spectral lines observed in the spectrum of an astronomical object with the rest wavelengths of the line. In the case of galaxies the absorption lines of the stellar spectra or emission and absorption lines of interstellar gas clouds can be used for this purpose. Since the spectral lines observed in galaxy spectra result from the superposed contributions of many stars and interstellar clouds, which are moving relative to the center of gravity of the galaxy, these lines have broad and complex profiles. To first approximation the orbital motions of the stars and the gas clouds result in a broadening of the lines. The resulting

Doppler widths are of the order of the orbital velocities of the stars and the gas clouds in the galaxy. These velocities may reach several 10^2 km s^{-1}. In addition, in star-forming galaxies stellar and interstellar absorption lines originating from resonance transitions often show systematic blueshifts of up to $\approx 10^3$ km s^{-1}. The blueshifts of high-ionization lines normally are due to the stellar winds of hot stars. The blue-shifts of the low-ionization lines are caused by galactic outflows of the interstellar gas. As a result of these line shifts, galaxy redshifts based on individual strong absorption lines tend to depend on the lines chosen, and they usually deviate systematically from the "systemic redshift," defined as the redshift of an ideal light source in the center of gravity of a galaxy. Stellar and galactic winds can also cause shifted emission lines if the emission is superposed on a shifted absorption component, if the winds are not isotropic, or if dust absorption obscures parts of isotropic outflows. Least affected by such effects are forbidden emission lines and stellar absorption lines originating from high atomic energy levels which are not populated in the thin gas of the stellar winds and the ISM. Therefore, redshifts derived from high-excitation photospheric lines and from forbidden interstellar emission lines are expected to provide a good approximations for the "systemic redshifts." However, forbidden emission lines are not always observed in galaxy spectra, and stellar absorption lines originating from high-energy levels often are too weak or (due to the broadening by the orbital motions) too heavily blended to be measurable in the low-S/N spectra of very distant galaxies.

Because of these problems with individual spectral lines, today redshifts are normally not derived by measuring individual lines. Instead precise redshifts of distant galaxies are usually derived by cross-correlating the observed spectra on a logarithmic wavelength scale with a template spectrum of the same type, for which the redshift is accurately known. If S_T is an unshifted template spectrum and S_{obs} an observed galaxy spectrum the cross-correlation function can be approximated by

$$C(z) \sim \int S_{obs}(\log \lambda)\, S_T(\log(\lambda + \lambda z))\mathrm{d} \log \lambda \tag{6.1}$$

where the integration extends over the observed spectral range. The redshift of the observed galaxy then follows from the position of the maximum of the function $C(z)$. Because of the integrations over the observed spectral range, the cross-correlation method makes optimal use of the spectral information. The error of the resulting redshift can be estimated from the width of the maximum of the function C. Moreover, by calculating cross-correlations with template spectra of different type, the amplitude of the maximum can be used to determine which type of template spectrum gives the best fit to the observed spectrum.

Like other data derived from spectra, the accuracy of redshift derivations depends on the spectral resolution R. However, the natural width of spectral lines in galaxy spectra limits the accuracy of redshift measurements of distant galaxies to about 10^{-4}. If individual lines are used, the error also depends on the sampling (which usually is defined by the detector pixel size). The cross-correlation technique is

somewhat less sensitive to the sampling since the positions of many spectral lines are averaged.

The cross-correlation function obviously measures the position of the spectral lines as well as the position of breaks and broad blends in the spectrum. Thus, this technique can be used even for spectra of very low resolution, where individual lines cannot be resolved.

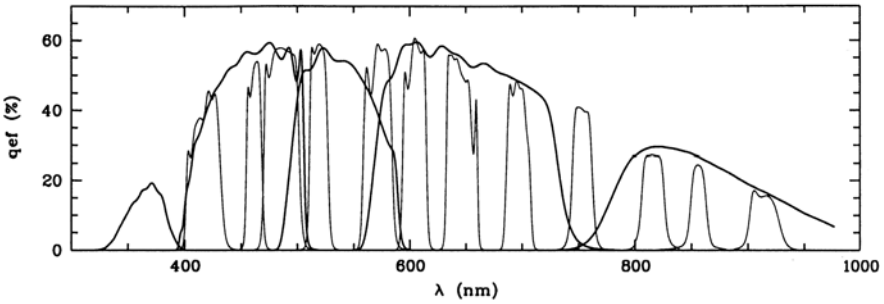

Fig. 6.1 Filter transmission functions for a photometric system (COMBO17) designed for deriving accurate photometric redshifts and galaxy types from photometric data. The system reaches a redshift accuracy of about 1% in $1 + z$. (From Wolf et al. [563])

As noted already in the context of identifying high-redshift galaxies in Sect. 5.1.2, breaks and broad blends in galaxy spectra can also be detected and measured by means of photometric observations, where these features cause characteristic differences in the flux measured at different wavelengths. With increasing redshift these features move through the different filter bands of a photometric systems. Therefore, the flux differences depend on the redshift. If the flux ratios are calibrated with galaxy spectra of known type and redshift, photometric observations can be used to derive redshifts. This "photometric redshift" technique was initially suggested (and applied for the first time) by W. A. Baum in 1961 [32]. With modern imaging detectors color indices and photometric redshifts can be obtained over relatively large fields simultaneously for all objects which are brighter than the sensitivity limit of the detector. The observable redshift range and the accuracy of the photometric redshifts depend on the widths and the number of filters used. With standard broad-band filters, covering the spectral range from the near-UV to the NIR, reliable photometric redshifts can be measured with mean errors of about $0.03\ (1 + z)$ (e.g., [189, 360]). Photometric systems using intermediate-band or narrow-band filters (such as the COMBO17 system, see [563] and Fig. 6.1) can reach accuracies of about 1% of $1 + z$ [39]. Computer codes for calculating redshifts from photometric data have been developed by many different groups. Some of these codes are publicly available (e.g., [57]).

Increasing the number of filters and narrowing the filter bands normally results in more accurate redshifts. On the other hand, the observing time increases with the number of filters, and fewer photons are collected in narrower bands. Therefore,

there is a trade-off between the accuracy and the time requirements (or the flux limit) of photometric redshifts.

Photometric redshifts can be derived from observations at all wavelengths where well-defined continuum breaks or broad spectral features are present in the spectra. An example outside the optical-IR range are photometric redshifts of dust-obscured galaxies. As described, e.g., in [244], for these objects photometric redshifts can be derived from FIR and radio observations, making use of the characteristic energy distribution of the emission of dusty starburst galaxies, as presented in Fig. 2.21.

6.2.2 Distances

Galaxies are too distant for geometric parallax measurements. Therefore, all direct distance derivations of galaxies are based measurements of the apparent brightness or the angular size of objects with known absolute luminosities ("standard candles") or of known linear sizes ("standard rods"). The most important "standard candles" for measuring large distances are the supernovae of type Ia (Sect. 2.1.5). Other potential high-redshift standard candles or standard rods are galaxies where the absolute luminosity or size can be derived by means of the Tully–Fisher relation, the Faber–Jackson relation, or the fundamental-plane relation (see Sect. 2.2.1).

All these methods involve the observed apparent brightness of objects with known intrinsic luminosity. Relations between the observed total flux from an object of luminosity L at a redshift z (or distance $a_0\, r$) in the expanding universe have been derived in Sect. 3.1.5 (see, e.g., Eq. 3.25).

At present, the methods for direct distance measurements listed above can be applied to galaxies with $z \leq 1.5$. With the next generation large telescopes with apertures $\geq$30 m it may be possible to extend these direct distance measurements to redshifts of $z \approx 4$. For more distant galaxies the redshift z has to be used to infer the distance. Relations between z and other distance measures again have been discussed and are given in Sect. 3.1.

In the local universe, where the cosmic expansion can be neglected, the distance is a well and uniquely defined quantity. On the other hand, on larger scales several different distance measures are in use. In Sect. 3.1.1 the "angular size distance" coordinate r and the "radial distance" coordinate r_D have been defined. Since the observational data indicate that the space curvature of our universe is zero (or at least very close to zero) in our actual cosmos r and r_D are (within the observational accuracy) equal.

In the astronomical literature often a third distance definition is used, which is called "luminosity distance." The luminosity distance D_l is defined by the relation

$$F = \frac{L}{4\pi\, D_l^2} \tag{6.2}$$

where L is the luminosity of a source, and F is the observed radiation flux. From a comparison of Eq. (6.2) and Eq. (3.25) we get for a flat cosmos

$$D_l = a_0 r(z)(1+z) = \frac{c(1+z)}{H_0} \int_0^z \frac{\mathrm{d}z}{E(z)}. \tag{6.3}$$

where $E(z)$ is given by Eq. (3.17) or (3.19). For $z \to 0$ the three distance measures $a_0\, r$, $a_0\, r_D$, and D_l become equal. For $z \to \infty$ D_l approaches infinity, while the angular-size distance and the radial distance remain finite.

6.3 Deriving Galaxy Luminosities

6.3.1 The Basic Steps

If the redshift and the observed spectrum of a distant galaxy is known, the intrinsic spectral flux and luminosity can be calculated from the observed flux using the relations given in Sect. 4.3.2. The resulting luminosity and SED correspond to the radiation which escapes from the galaxies into the intergalactic space. For a comparison with models often the intrinsic luminosity and spectral energy distribution of the total light emitted by the stars has to be known. In order to derive these quantities, the observed spectral flux has to be corrected for the absorption and emission of the interstellar matter of the observed galaxies. Moreover, lensed galaxies have to be identified and the gravitational flux amplification has to be removed.

Most important in this context is an accurate correction for the internal dust absorption. Since the absorption cross section of interstellar dust grains increases with decreasing wavelength, dust absorption is particularly strong in the rest-frame ultraviolet, which is the best studied spectral range of the high-redshift galaxies.

As discussed in Chap 7, the most distant known galaxies are starburst galaxies. In the local universe stars normally form from dense, cool interstellar matter. As a result the internal dust absorption of galaxies is correlated with the star formation rates, and dust absorption tends to be particularly strong in galaxies with high SFRs. While dust absorption in elliptical galaxies often is negligible, it tends to be exceptionally high in the case of massive starburst galaxies, where sometimes all the UV and visual light is completely absorbed by the dust (see Sect. 7.1). Because of these facts, dust absorption is an important issue for high-redshift galaxies.

6.3.2 Corrections for Dust Absorption

As discussed in Sect. 2.2.5 the absorption of short-wavelength radiation by dust grains results in a heating of the dust. Since the dust re-radiates the observed energy at FIR wavelengths (shifted to the submm range for high-redshift galaxies), the total (bolometric) luminosity remains essentially unchanged. But, because of the varying atmospheric transmission and the different sensitivities of instruments at different wavelengths, observing the full spectrum extending from the (observer's frame) visual to the millimeter range normally is not feasible. Often only optical

data are available. In this case the dust absorption has to be determined, and the spectral flux has to be corrected in order to obtain the spectrum and the luminosity of the stellar populations.

In principle the dust absorption can be derived by comparing the observed line spectra and the continua with synthetic spectra (described in Sect. 6.4) of dust-free galaxies. Because of the decrease of the dust absorption with increasing wavelength, dust absorption results in a "reddening" of the observed continuum while the line spectrum of the stellar population remains unchanged. Thus, a detailed comparison of the SED and the line spectrum can be used to estimate the dust absorption. However, a reliable separation of the dust reddening effects from the effects of the stellar age distribution and of the chemical composition requires spectra of relatively high spectral resolution and a good S/N, which cannot be obtained for faint high-redshift galaxies. Therefore, normally other methods are used to derive the dust reddening.

Most of these methods make use of the almost uniform slope of the UV continua of unreddened young starburst galaxies. As shown by Fig. 6.5, the slope of the (rest-frame) UV continuum between 130 and 180 nm of a starburst changes only slightly during the first ~ 10 Myr. Quantitatively, the continuum slope in this wavelength range is often characterized by the parameter β, which is derived (and defined) by fitting a function

$$F_\lambda \propto \lambda^\beta \tag{6.4}$$

to the observed spectral flux in the wavelength interval defined above. Over a wide range of starburst parameters unreddened young starbursts have β in the range $-2.6 < \beta < -2.2$ (e.g., [312]). For most dust-free starburst galaxies $\beta = -2.5$ is a good approximation. For local starburst galaxies β was found to correlate well with the dust absorption derived by comparing the theoretical Balmer decrement with the strengths of observed Balmer emission lines of starburst galaxies [84]. This correlation (and other independent derivations of the amount of dust absorption) provides a calibration of β as a measure of the light attenuation by the dust. For young starburst galaxies, where the intrinsic β can be predicted, the difference between the observed and the expected UV continuum slope can be used to estimate directly the amount of internal dust reddening and dust absorption [84, 85, 311, 351, 471]. With the calibration mentioned above we get for local starburst galaxies

$$A_{150} \approx 2.19(\beta - \beta_0) \tag{6.5}$$

where A_{150} is the flux attenuation (in magnitudes) at $\lambda = 150$ nm, β is the observed UV continuum slope parameter, and $\beta_0 \approx -2.5$ is the continuum slope which is observed in the absence of dust absorption. The exact value of β_0 depends on the star formation history and the chemical composition. These two parameters can, in principle, be inferred from the line spectra.

Equation (6.5) is referred to as the "Calzetti law." It provides a good fit to the data of local starburst galaxies, as well as for many distant star-forming systems.

Therefore, the Calzetti law is widely used to correct the luminosities and spectra of distant starburst galaxies for internal dust extinction. However, the results of computer simulations of the radiative transfer in dusty starburst galaxies (reproduced in Fig. 6.2) indicate that differences in the dust properties and of the dust distribution can significantly affect the relation between the light attenuation and β. A careful analysis of the spectra of high-redshift starburst galaxies indicates that such differences of the dust properties between individual starburst galaxies are actually observed [378, 379]. As a result, the reddening corrections of high-redshift starburst galaxies introduce uncertainties into the derivation of the total luminosities of these objects.

6.3.3 Luminosity Functions

Apart from obtaining the luminosities of individual galaxies, an important objective of survey programs is the derivation of the luminosity distribution or "luminosity functions" (LF, see Sect. 2.2.2) of the distant galaxies. Obviously, LFs can be constructed, if all galaxies in a known volume are observed down to a fixed limiting absolute luminosity. While a complete count of the bright galaxies normally does

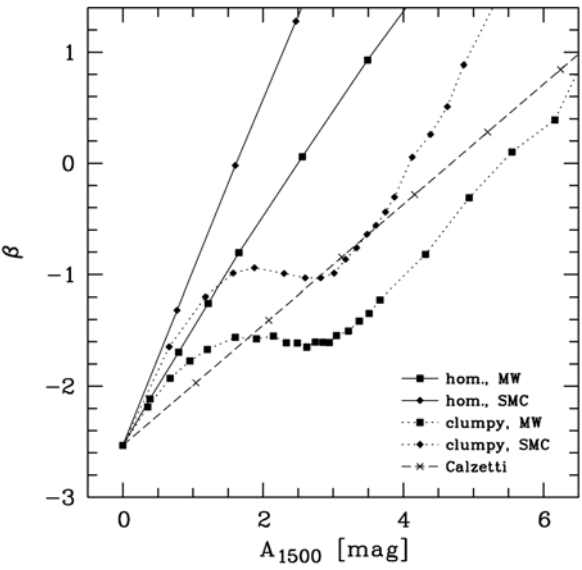

Fig. 6.2 The relation between the UV continuum slope β and the light attenuation at 1500 Å predicted from model computations for different properties of the interstellar dust in starburst galaxies. The *crosses* and the *dashed line* represent the Calzetti law observed in UV-bright local starburst galaxies. The other curves give the relation between β and the light attenuation assuming the dust properties observed in the Milky Way galaxy (MW) and the Small Magellanic Cloud (SMC) galaxy. In both cases the models were calculated for a homogenous dust distribution and for a clumpy distribution of the dust. (From Noll and Pierini [378], where details are described)

not pose a problem, there is always a danger of missing faint and small galaxies. Due to S/N problems faint galaxies tend to be missed by extractor codes. Moreover, they may appear to be parts of superposed foreground (or background) galaxies. Therefore, for deriving LFs the completeness limits of surveys have to be carefully determined. This can be done, e.g., by artificially adding faint galaxy images to the observed sky frames and by studying the success rate of retrieving the additional galaxies as a function of their properties.

For deriving LFs covering a large range of luminosities, often large and small fields at different locations have to be combined. Because of the small-scale inhomogeneity of the present-day universe, this may introduce significant errors into the statistics and the resulting LFs. To avoid such errors, methods have to be applied which measure the distributions in the individual fields rather than absolute counts in the different fields. A technique commonly used for this purpose is the Sandage-Tammann-Yahil (STY) method [448]. But, depending on the specific properties of the observational data, other methods can give superior results (e.g., [64, 153]).

6.4 Stellar Populations and Ages

Since high-redshift galaxies cannot be resolved into individual stars, information on their stellar content has to be derived from the observed integrated spectra of these objects. Examples of observed spectra representing galaxies of different stellar populations are reproduced in Fig. 6.3. As shown by this figure, on the basis of such spectra star-forming galaxies can be clearly identified and separated from galaxies with passively evolving stellar population.

In order to derive quantitative information on the stellar populations, the observed spectra usually are compared to "synthetic" galaxy spectra. The synthetic spectra are constructed by calculating sums of observed or computed stellar spectra of different types, assuming a plausible distributions of the (initial) stellar masses of the individual stars, and making assumptions on the stellar ages and age distributions. Computer codes producing synthetic galaxy spectra have been developed by many different groups (e.g., [78, 102, 173, 207, 310, 312, 336, 439, 543]). Many of these codes and extensive digital libraries of synthetic spectra are publicly available and can be retrieved through the Internet. (For details see the publications cited.) The codes can be used to predict the spectra, luminosities, and color indices of stellar systems with known properties as a function of age. By varying the input parameters until the synthetic spectra match the observed spectra it is—at least in principle—possible to constrain or to derive the physical parameters of the observed stellar populations.

The numerical codes for calculating synthetic spectra naturally contain many free parameters. Among these parameters are the total mass of the stars included, the initial stellar mass distribution (IMF), the star formation rates as a function of time (usually referred to as the "star formation history" or "SFH"), the chemical composition of the star-forming gas, and the mass exchange (and other interactions)

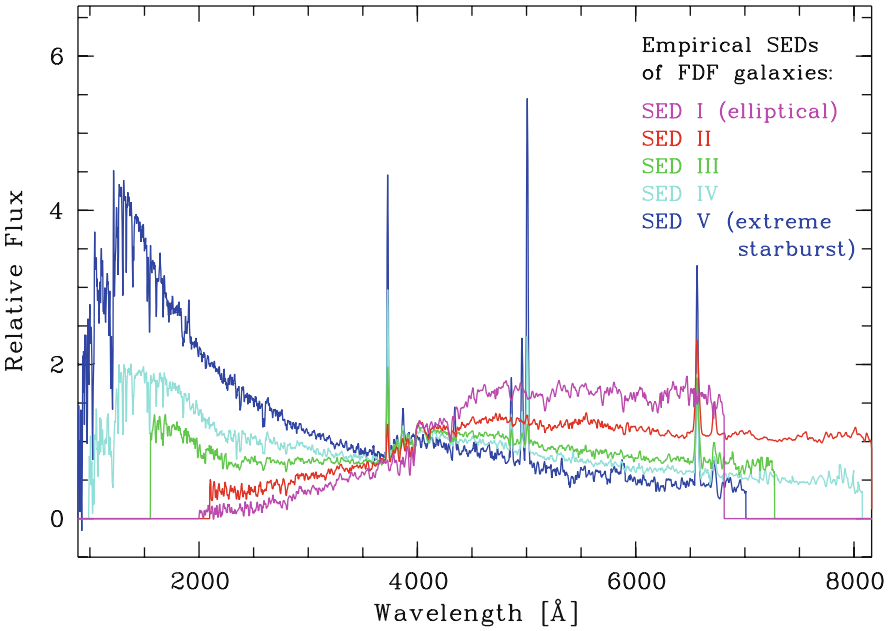

Fig. 6.3 Rest-frame composite spectra of distant galaxies with different star formation activity. All spectra are taken from the FORS Deep Field spectroscopic survey. The specific star formation rates range between practically zero in the elliptical galaxies (SED type I) to extreme starburst activity (SED type V). The spectra have been generated by averaging the individual spectra of the corresponding samples. They have been normalized to have equal mean flux in the wavelength range of 3500–4500 Å. (Courtesy S. Noll)

with other galaxies and the IGM. For the distribution of the stellar masses usually a Salpeter-type initial mass function (Eq. 2.1) is assumed. Different slope parameters and different upper and lower mass cut-offs of the Salpeter IMF are used to simulate the influence of variations of the assumed IMF. For the stellar ages often two limiting cases are calculated. The first case assumes a single "instantaneous" burst of star formation, after which no further formation of stars takes place. The second case assumes a "continuous star formation" at a constant star formation rate. In the case of continuous star formation, stars of all masses contribute to the combined spectrum. In the case of a single instantaneous starburst, initially the bright but massive stars dominate, while later on (when the massive stars have ended their short lives as SNe) the spectra will be due to the intermediate-mass and low-mass stars only.

Examples of synthetic spectra (calculated with the STARBURST99 code of [312] for an instantaneous starburst) are given in Figs. 6.5 and 6.4. As shown by these figures, with the parameters assumed, during the first ≈ 2 Myr the spectrum of an instantaneous starburst is dominated by the FUV radiation of hot O-type stars. At high metallicities, there is a short period (of about 3 Myr) where (due to the formation and brief existence of very hot WR stars) the spectrum extends to even

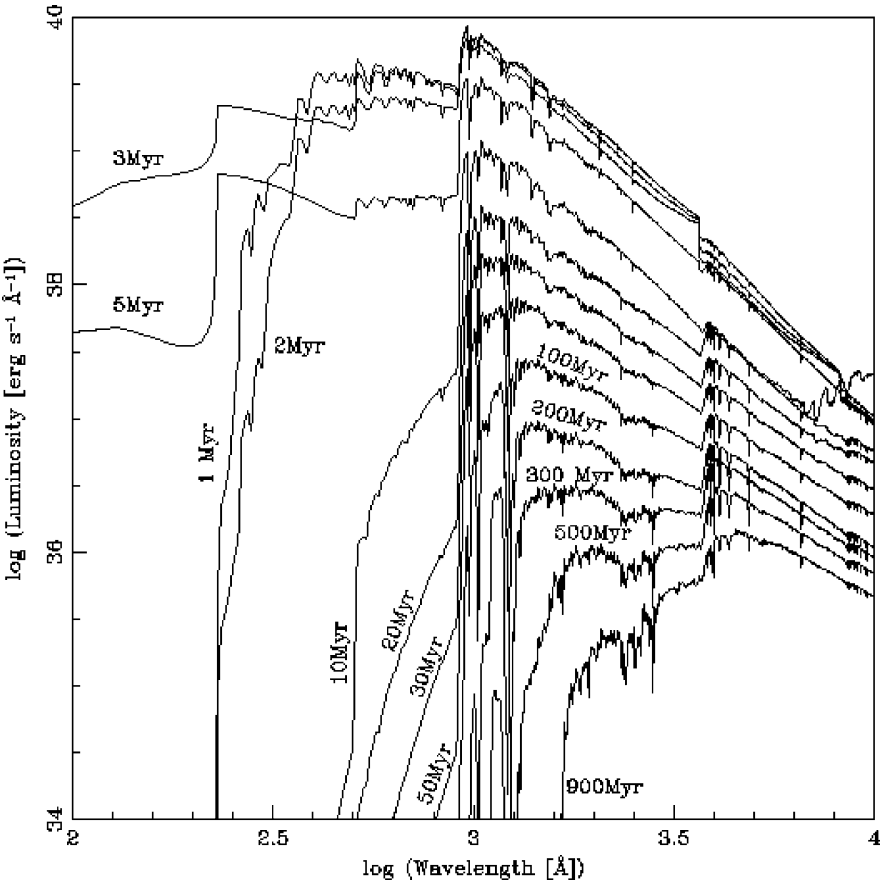

Fig. 6.4 Example of synthetic galaxy spectra assuming stellar populations which are formed in an instantaneous starburst with a solar initial chemical composition ($Z = 0.02$). The figure shows the spectra as they are expected to appear at population ages between 1 Myr and 900 Myr. (From Leitherer et al. [312])

shorter (EUV) wavelengths. After 10 Myr the EUV flux drops dramatically due to the Lyman continuum absorption in the atmospheres of the now dominating A and B stars. At ages of > 1 Gyr the spectrum is essentially due to the visual and IR radiation of cool, low-mass ($M \leq 2\,M_{\odot}$) stars.

While the FUV and EUV radiation field is important for the internal properties of the star-forming galaxies, this part of the spectrum normally cannot be directly observed in high-redshift galaxies, since the Lyman continuum is efficiently absorbed by neutral hydrogen atoms in and around the galaxies (e.g., [472]). Therefore, a direct comparison of synthetic and observed galaxy spectra is essentially restricted to the spectral range longward of the Lyman limit.

Figures 6.5 and 6.4 also show that old stellar populations with ages >1 Gyr radiate progressively more light in the red and NIR spectral range. For high-redshift

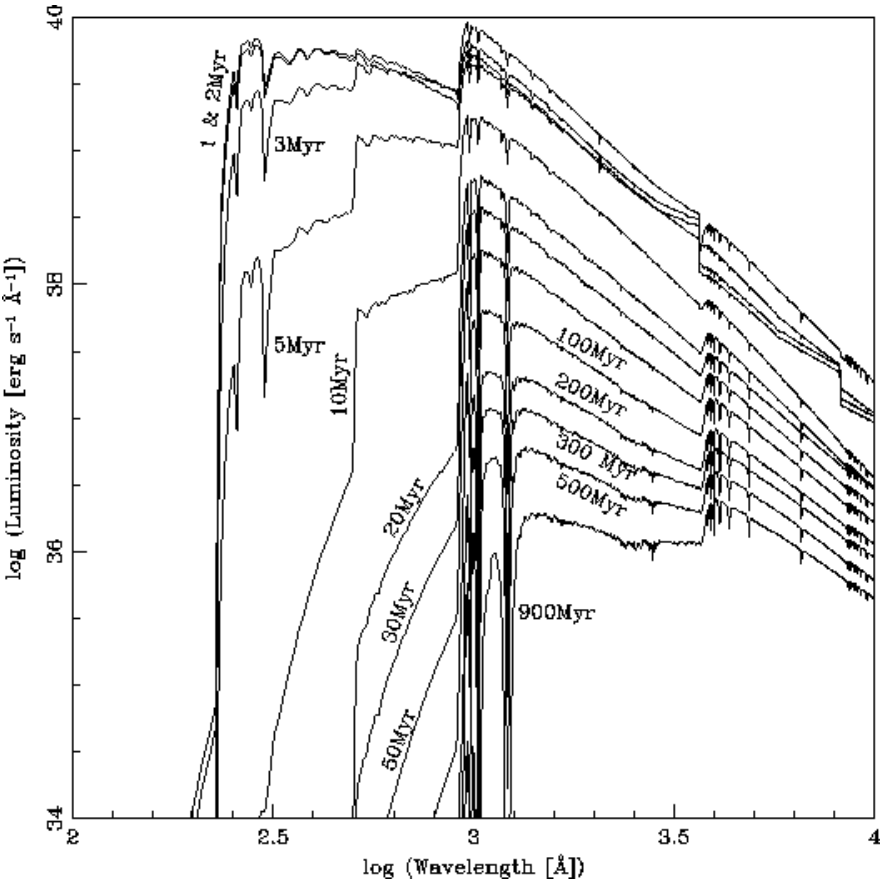

Fig. 6.5 Same as Fig. 6.5, but assuming a lower abundance of heavy elements ($Z = .001$). (From Leitherer et al. [312])

galaxies this spectral range is shifted into the mid-infrared, which cannot be observed well from ground-based observatories. Therefore, these figures also illustrate the great importance of space-based MIR observations for population studies of high-redshift galaxies with evolved populations.

From Fig. 6.5 it is clear that a comparison with synthetic spectra provides a powerful tool for estimating populations and stellar ages of galaxies. However, because of the many free parameters, an accurate quantitative separation of the effects of age, IMF, and chemical composition is not always easy and normally requires spectra of relatively high resolution. A discussion of the potential, the limitations, and of some of the shortfalls of the spectral-synthesis technique can be found in [431].

Since spectral synthesis models are based on *stellar* spectra only, they normally do not include spectral features due to the internal interstellar medium of the galaxies. (An exception are the continua of the interstellar gas emission, which are

included in some synthetic starburst spectra.) Therefore, when comparing observed and synthetic spectra, the internal absorption and emission by the interstellar dust and by of lines and continua by interstellar atoms, ions, and molecules have to be taken into account during the analysis (see also Sects. 6.3 and 6.8).

6.5 Mass Estimates

Synthetic spectra use among their input parameters the total mass of the stars involved, and they predict the resulting luminosity as a function of wavelength. Therefore, by comparing observed and synthetic spectra the mass present in the form of stars in a galaxy can be estimated.[1] A detailed description and discussion of the mass derivation by comparison with synthetic spectra is given, e.g., in [147].

Spectral synthesis models also provide mass-luminosity (M/L) ratios for each computed spectrum. However, the luminosity of individual stars depends strongly on their mass as well as on their evolutionary state. Thus the M/L of a galaxy depends sensitively on the star formation history. To obtain reliable mass estimates either the SFH has to be constrained very well by the spectra, or suitable strategies have to be used to minimize the errors resulting from inaccurately estimated star formation histories (e.g., [58]).

As noted above, the observed luminosities provide information only on a galaxy's mass present in the form of stars. To derive the total mass (including the interstellar gas and the dark matter) the internal kinematics of the galaxy has to be studied by means of spatially resolved spectroscopy. Because of the small angular size of high-redshift galaxies, this usually requires integral field spectroscopy techniques in combination with adaptive optics systems. Details and some results are discussed in Sect. 7.1.2.

6.6 Star-Formation Rates

Since the star-formation rate (SFR) is among the input parameters of the codes producing synthetic spectra, a comparison between observed and synthetic galaxy spectra also yields the SFRs of observed galaxies. However, as with other parameters, the accuracy of SFRs resulting from spectral comparisons rests on a correct separation of the different effects which determine the spectral details. Moreover, often reasonably accurate estimates of the SFRs can be derived by more direct methods. These methods are based on the fact that in star-forming stellar systems the UV continuum is produced almost entirely by the newly formed and shortlived

[1] In the extragalactic literature the total mass present in a galaxy in the form of stars is referred to as the "stellar mass" (to distinguish it from the "dynamical mass," which includes the interstellar gas and the dark matter). In other fields of astrophysics the term "stellar mass" is used for the mass of individual stars. Therefore, the meaning of "stellar mass" has to be inferred from the context.

massive stars, while the red and NIR regions of the spectra are often dominated by old stars. Therefore, blue-to-red flux ratios or color indices (such as $U - V$) provide information on the SFR (see, e.g., Fig. 2 of [265]). Although these estimates are not very accurate, color indices can provide a first approximation of the SFR of an observed galaxy.

Since practically all the UV continuum of galaxies is due to recently formed stars, the emitted UV flux is approximately proportional to the star formation rate. Moreover such a proportionality can also be expected for the strength of the recombination-line radiation of the interstellar gas, which is ionized by the EUV continuum. Because of internal absorption, the EUV ($\lambda < 912$ Å) continuum cannot be observed directly. But the adjacent FUV continuum of the high-redshift galaxies is shifted to the visual spectral range, where it can be observed with the large ground-based optical telescopes. Conveniently, in young starburst galaxies the spectral flux F_ν and the spectral luminosity L_ν in the wavelength range $1500 < \lambda < 2800$ Å are practically independent of the wavelength. According to the calibration given in [265], assuming a Salpeter IMF and continuous star formation over a period of $\geq 10^8$ yr, the SFR of a star-forming galaxy is proportional to this constant intrinsic spectral luminosity L_ν, and is given with good approximation by the simple relation

$$\mathrm{SFR} = 1.4 \times 10^{-21} \frac{L_\nu}{\mathrm{W\,Hz^{-1}}} \mathrm{M_\odot\,yr^{-1}}. \tag{6.6}$$

Note that the numerical factor is correct only for the assumption given above. For younger stellar populations the SFR per unit UV luminosity is lower, and for older systems higher values of the ratio SFR/L are found. (For details see, e.g., [265].)

As pointed out in Sect. 5.1.3, in starburst galaxies containing a significant amount of neutral hydrogen, most or all Lyman continuum photons tend to become converted into Lyα photons. Therefore, the SFR can also be derived directly from the Lyα luminosity. Following again the calibration given in [265] (based on the same assumptions), the corresponding relation becomes:

$$\mathrm{SFR} = 9.1 \times 10^{-36} \frac{L_{\mathrm{Ly}\alpha}}{\mathrm{W}} \mathrm{M_\odot\,yr^{-1}}. \tag{6.7}$$

This relation is usually applied to high-redshift galaxies where the (rest-frame) UV continuum is too weak to be measurable, while Lyα emission flux (which is less affected by the sky background) can still be evaluated (see Sect. 7.1). Equation (6.7) is correct only if all Lyman continuum photons are converted into Lyα photons *and* if all Lyα photons escape the galaxy. But because of their high resonance-scattering cross-section the Lyα photons are scattered in the interstellar medium. Therefore, their geometric path in the intrinsic ISM is longer than that of continuum photons, and their probability to be absorbed by dust grains is much higher. Even a small amount of dust can completely quench the Lyα emission. Because of this effect Eq. (6.7) tends to underestimate SFRs. But the Lyα flux can at least provide a lower limit of the SFR.

Since the theory of photoionized gases accurately predict the ratios between the strengths of all hydrogen lines produced in the recombination process (e.g., [142]), relations corresponding to Eq. (6.7) can be readily formulated for other H I transitions. For the strong Balmer line Hα we get (again with the parameters noted above):

$$\mathrm{SFR} = 7.9 \times 10^{-35} \frac{L_{H\alpha}}{\mathrm{W}} \mathrm{M}_\odot\,\mathrm{yr}^{-1}. \tag{6.8}$$

This relation again assumes that the Lyman continuum photons cannot escape from the galaxy and are all converted to Lyα photons. Resonance scattering from the $n = 2$ level of the hydrogen atom is negligible in the ISM, and at the longer wavelength of Hα dust absorption is less efficient than at Lyα. Therefore, the Hα photons usually freely escape from a starburst galaxy. As a result, compared to Eq. (6.7) the relation (6.8) often provides a more accurate and more reliable estimate of the SFR.

Even less affected by dust are the IR and radio recombination lines. But in high-redshift galaxies these lines tend to be too weak or redshifted to wavelengths which cannot be observed easily from the ground.

For galaxies where all the UV starlight is absorbed by dust and re-emitted at far infrared wavelengths, and where stars or AGN do not contribute significantly to the FIR emission, the total IR emission can be used to estimate SFRs using the relation

$$\mathrm{SFR} = 4.5 \times 10^{-37} \frac{L_{FIR}}{\mathrm{W}} \mathrm{M}_\odot\,\mathrm{yr}^{-1}. \tag{6.9}$$

where L_{FIR} is the total IR luminosity of the dust, integrated over the wavelength range of 10 μm–1 mm. The calibration of the relation (6.9) again is based on [265]. The reliability of SFRs derived from the IR continuum depends on how well the conditions listed above are fulfilled, and how well they can be verified.

Observations of H II regions in the local universe show that the flux ratios between the Hα line and certain collisionally excited forbidden lines or blends show only a relatively small range of variations. Among these lines is the strong [O II] 3727 Å doublet ([192], [265]). Thus, for galaxies where (e.g., because of an inconvenient redshift) Hα cannot be observed, SFRs can be derived by observing the [O II] line (or another forbidden line). Using the known ratio between the forbidden line and Hα, the Hα strength can be estimated and the SFR can then be determined from Eq. (6.8). Naturally, such indirect SFR estimates are less accurate than those based directly on the Hα flux. (For details, see [265].)

Since the frequent supernovae of young stellar populations produce radio-continuum emission and X-rays, in principle, SFRs can also be derived from the radio flux and the X-ray flux of starburst galaxies. In fact, the radio continuum flux is used to find candidates for dust-obscured massive submillimeter galaxies. However, the radio and X-ray emission of most of the $z > 2$ starburst galaxies is below the sensitivity limit of the present-day radio and X-ray telescopes. But with

future instruments such observations may well become feasible. The flux at these wavelengths may then provide another independent method for deriving SFRs.

6.7 Deriving the Chemical Composition

For nearby galaxies the chemical composition can be derived by means of several different methods. Firstly, the relative abundance of the chemical elements in the atmospheric layers of individual stars can be determined by analyzing their spectra. If stars are selected where the composition of the atmospheric layers has not (yet) been modified by the nuclear reactions in the stellar interior, by mass accretion, or by mass loss, the stellar spectra also provide information on the chemical composition of the interstellar gas from which the stars were formed. More directly the chemical composition of the interstellar gas can be determined from an analysis of the absorption and emission lines of the interstellar gas.

Analyzing the spectra of individual stars obviously is not an option in the case of high-redshift galaxies. Deriving chemical abundances in high-redshift galaxies from interstellar absorption lines is possible in principle. But in interstellar space, atoms and ions are normally in their energetic ground state. Thus, interstellar absorption lines typically correspond to resonance transitions, which tend to saturate at relatively small column densities. As a result, in the integrated spectra of distant galaxies the strength of these lines usually depend more on the internal gas kinematics than on the chemical abundances. Moreover, detailed observations show that the cloudy structure of the interstellar medium has a major effect on the strength of the interstellar absorption lines in high-redshift galaxies (see Sect. 7.1.2). Disentangling these different effects require spectra of high spectral resolution, which normally cannot be obtained for the faint high-redshift galaxies.

Less problematic is the analysis of the interstellar emission lines of high-redshift galaxies. These lines are produced in gas which is ionized by hot stars. Therefore, the line spectrum emitted by the corresponding H II regions (Sect. 2.2.5) can be predicted with reasonable accuracy from photoionization models (For details and a review of the extensive literature on this subject see, e.g., [142] and [390].)

6.7.1 Abundances from H II Emission-Line Spectra

From photoionization models various emission-line ratios have been identified which can be used to derive the chemical abundance of the observed elements relative to hydrogen and helium. In general the strength of the interstellar emission lines depend on the element's relative abundance, on the radiation field, and on the density and temperature of the gas. If a sufficient number of different lines are observed, the effects of density and temperature can be eliminated and the abundance can be well constrained.

The effects of the radiation field and the temperature on abundance derivations can be determined and eliminated most easily for elements where lines of several different ionization stages can be observed simultaneously. This is the case, e.g., for oxygen, which (in addition to being one of the most abundant elements) in H II region spectra has strong emission lines of O, O^+, and O^{++}. Moreover, on spectra of sufficient S/N, the gas electron temperature can be determined from the intensity ratios between the transitions from the first and the second excited level to the ground state of [O III] (e.g., [390]). Therefore, oxygen is the most frequently measured element in H II regions, and the heavy element abundance of H II regions is often characterized by the oxygen content. In many papers it is implicitly assumed that the abundance of other elements scales with the oxygen abundance. If this assumption is justified, the oxygen abundances derived by means of ionization models and accurately determined electron temperatures can be used to calibrate other abundance-sensitive line ratios of the H II region emission spectra.

A particularly valuable line ratio for measuring the oxygen abundance in H II regions was suggested by B. E. J. Pagel and his coworkers [399]. This line ratio, which is called R_{23}, is derived by dividing the total flux in the three oxygen lines or blends [O II] 3727 Å, [O III] 4959 Å, and [O III] 5007 Å, by the flux in the Hβ emission line. Thus, R_{23} is defined as

$$R_{23} = \frac{F_{3727} + F_{4959} + F_{5007}}{F_{H\beta}}. \tag{6.10}$$

Since the four lines involved are among the strongest emission lines of H II regions, the R_{23} ratio can be measured even in the spectra of rather faint objects. However, although very sensitive to the oxygen abundance, R_{23} it is not a monotonic function of the oxygen content. Starting from a small oxygen abundance R_{23} increases with the oxygen-to-hydrogen ratio, reaching a maximum near O/H $\approx 3\times10^{-4}$. For higher values of O/H the ratio R_{23} decreases again.[2] Therefore, except at the maximum, a given ratio R_{23} matches two different O/H values, one corresponding to a high and the other corresponding to a low oxygen abundance. Selecting the correct one of these two value requires additional line ratios which (because of lower line strengths or blending) often need higher resolution spectra. (For details and a calibration of the R_{23} ratio see, e.g., [142, 270, 367, 390].)

If spectra of sufficient resolution and signal-to-noise ratio can be obtained, many other abundance-sensitive emission line ratios can be evaluated. Examples of commonly used line ratios are given in Fig. 6.6. The calibrations presented in Fig. 6.6

[2] That the relative oxygen line strength can decrease with an increasing oxygen abundance appears surprising on a first glance. But the decrease is readily understood by the fact that the increasing abundance of oxygen (and of other heavy elements) results in a stronger cooling of the gas by IR emission lines. Since the [O II] and [O III] lines are collisionally exited, they become weaker with the decreasing temperature, while the recombination line Hβ is little affected by the temperature decrease.

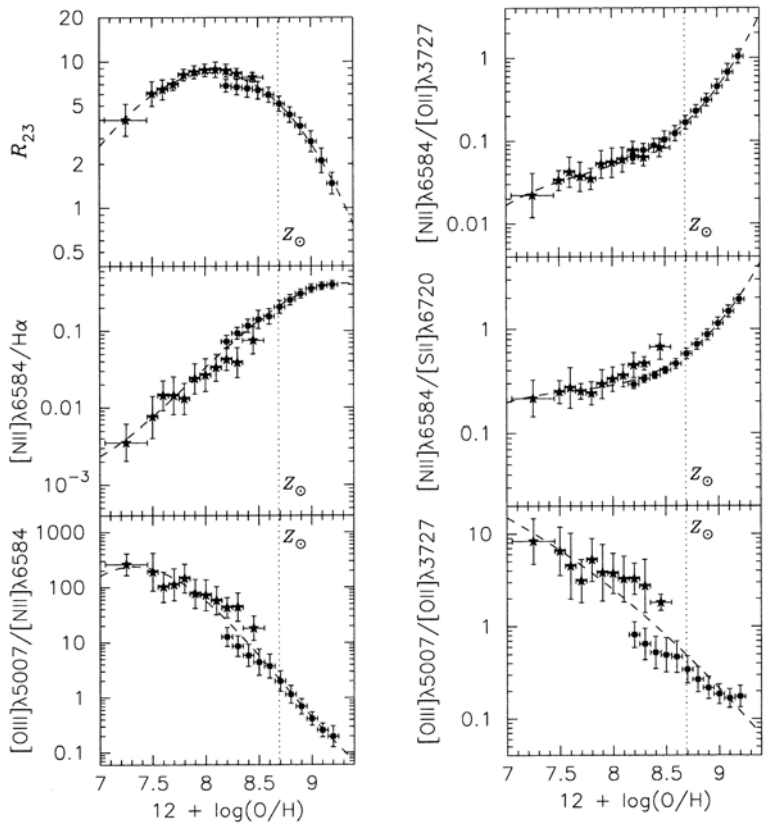

Fig. 6.6 Flux ratios of emission lines observed in galaxies as a function of the oxygen abundance. The *vertical dotted line* indicates the solar abundance. (From [367])

(from Nagao et al. [367]) are based on more than 50 000 galaxy spectra with well-determined oxygen abundances.

For high-redshift galaxies the spectral lines used by the R_{23} method (and many other suitable optical diagnostic line ratios) are shifted to NIR wavelengths. Hence, with ground-based telescopes the use of these line ratios is restricted to redshifts where the corresponding lines coincide with atmospheric spectral windows (i.e., the H, J, or K bands). Therefore, the diagnostic line ratio [Ne III] 3869 Å/[O II] 3727 Å (not included in Fig. 6.6) is of particular importance for studying high-redshift galaxies. Because of the short wavelengths of the lines involved, the [Ne III]/[O II] ratio can still be applied at redshifts $z > 3$. A calibration of this line ratio in terms of the oxygen abundance can again be found in [367].

In spite of the restricted wavelength range, the analysis of the H II emission lines is still the most important source of information on the chemical composition of galaxies. This method has the obvious advantage that the measurements can

be directly compared to observations in the local universe, and that the physics of photoionized nebulae is relatively well understood.

6.7.2 Abundances from the Integrated Stellar Spectra

While the chemical abundance of individual stars cannot be measured in high-redshift galaxies, some information on the chemical composition can be derived from the spectra of the observed integrated starlight. This is demonstrated by a comparison of Figs.6.4 and 6.5. These two figures contain synthetic spectra calculated with the same assumptions, except for a different heavy-element abundance. Among the most conspicuous differences in the resulting spectra is the absence of the FUV flux of WR stars at ages 3–5 Myr in the low-abundance case. At these ages, and in this wavelength range, the "low-metallicity" spectrum appears redder than the solar-abundance spectrum. On the other hand, with increasing age the absorption lines of the heavier elements reduce the flux in the near UV spectral range. As shown by the comparison of the two figures this "line blanketing" becomes particularly important for ages >100 Myr. Because of the reduced line absorption, the low-metallicity models now appear bluer than the solar-abundance ones. However,

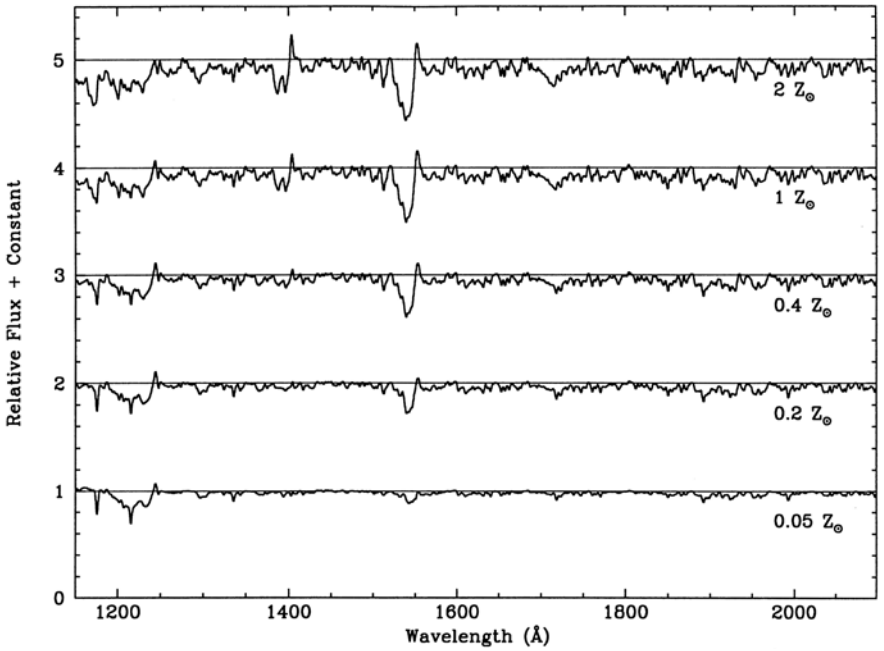

Fig. 6.7 Synthetic spectra of starburst galaxies showing the effect of chemical abundance differences on the strength of the photospheric UV absorption lines. The observed abundances range between twice the solar abundance ($2Z_{\odot}$) and 5% of the solar value (0.05 $Z_{\odot}$). (From Rix et al. [439])

the figures also show that the continuum colors depend strongly on the stellar ages. Therefore, chemical composition estimates based on the continua tend to be ambiguous and observed energy distributions often can be matched equally well with different combinations of age and chemical composition. In the literature this effect is referred to as the "age-metallicity degeneracy."

More reliable abundance information can be derived from the absorption line spectra of the integrated starlight. As shown by the examples of Fig. 6.7, in the starburst spectra the absorption lines of heavy elements become progressively stronger with increasing chemical abundances.

In the synthetic starburst spectra reproduced in Fig. 6.7, the abundance effect is most conspicuous for the strong resonance lines and blends, such as N V 1241 Å, Si IV 1400 Å, and C IV 1550 Å. A clear correlation between the absorption line strength and the chemical abundance is also found in the observed spectra of local starburst galaxies [225]. However, the observed line strengths show a significantly larger scatter around the correlation between the absorption strength and the chemical abundance than expected from the synthetic spectra. A reason of this larger scatter in the empirical data are interstellar components of the resonance lines. As pointed out above, in integrated galaxy spectra the strength of the interstellar lines depend more on the velocity field and the covering factor of the observed gas and less on the chemical abundance. As a result pure interstellar absorption lines of local starburst galaxies show only a weak correlation with the oxygen abundance derived from emission line ratios. On the other hand, as shown by Heckman et al. [225] the C IV and Si IV absorption line strengths of local starburst galaxies are well correlated with the oxygen abundance. This seems to be due to the fact that in starburst galaxies the C IV and Si IV resonance lines are dominated by stellar-wind components, which are very sensitive to the heavy-element abundance. A more detailed analysis of the data for the local starburst galaxies shows that the correlation is most clearly defined for the C IV resonance doublet.

Using the spectral data on local starburst galaxies collected by Heckman et al. [225], Mehlert et al. ([345]) derived an empirical calibration of the C IV absorption line strengths in terms of the fractional heavy element abundance Z. According to these authors we have

$$\log(Z/Z_{\odot}) = 0.13(\pm 0.02) \times W_0(CIV) - 1.10(\pm 0.12) \qquad (6.11)$$

where Z_0 is the solar abundance and $W_0(CIV)$ is the total rest-frame equivalent widths of the C IV resonance doublet. Because of the complex structure of the starburst spectra near this doublet, the exact calibration depends on the spectral resolution and on the details of the equivalent width measurements (see, e.g, [116]).

For the (so far relatively few) galaxies where C IV–based abundances and estimates based on other methods are available, the results have been found to be reasonably consistent. However, because of the scatter in the empirical relation between the C IV strength and the chemical abundance, C IV–based abundance estimates may not be reliable for individual galaxies. On the other hand, the C IV method is expected to work well for statistical studies, where it is particularly valuable, since

the strong C IV feature can be readily measured even in the spectra of rather faint galaxies. Compared to the optical H II emission lines, the C IV resonance doublet has the advantage of being observable at much higher redshifts.

While the strengthening of the resonance lines with increasing metallicity is the most conspicuous effect visible in Fig. 6.7, the weaker photospheric lines show the same trend. As these weaker lines normally correspond to transitions from excited states of atoms or ions, they have no interstellar contributions. Therefore, these lines are expected to correlate more tightly with the chemical composition than found for the resonance lines. As pointed out by S. A. Rix and coworkers [439], relatively robust measures of the chemical composition are the equivalent widths of three complex blends covering the wavelength intervals 1360–1380 Å ("1370 blend"), 1415–1435 Å ("1425 blend"), and 1900–2050 Å ("1978 blend"). In Fig. 6.8 equivalent widths of these blends, as measured in the synthetic spectra of [439], are plotted as a function of the metallicity. Also included in the figure are the equivalent widths of the C IV resonance doublet measured in the same synthetic spectra. The figure shows that all plotted equivalent widths increase monotonically with increasing heavy element abundance. For the 1370, 1425, and 1978 blends the

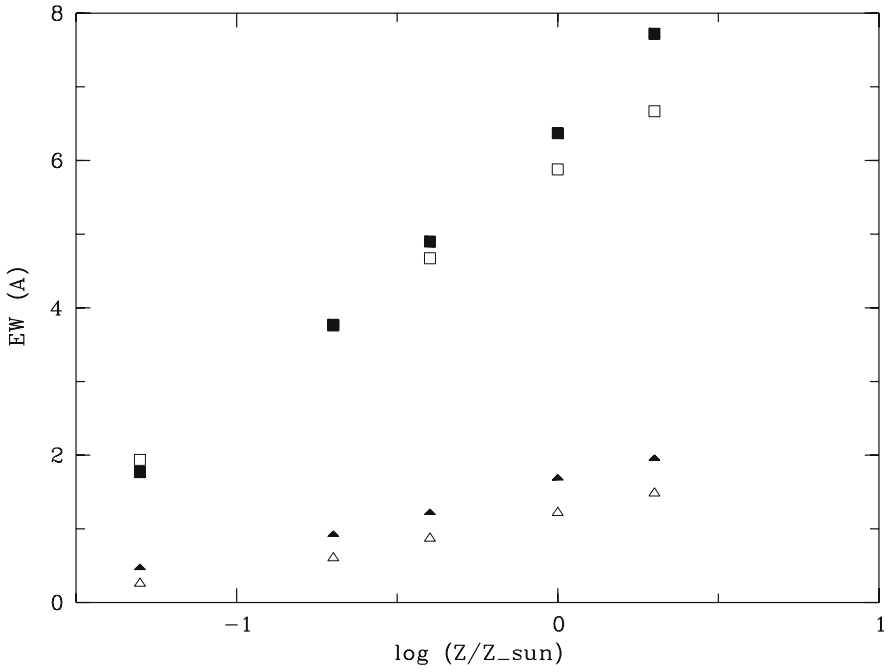

Fig. 6.8 Predicted equivalent widths (in Å, calculated using the synthetic spectra of [439]) as a function of the heavy element abundance Z for the 1370Å blend (*filled triangles*), the 1425Å blend (*open triangles*), the 1978Å blend (*open squares*), and the C IV resonance doublet (*filled squares*) in young (≈ 100 Myr) starburst galaxies with a Salpeter IMF. (Based on [347] and [439], where details are described)

correlations in Fig. 6.8 can be directly used for estimating chemical abundances from observed equivalent widths. But, because of the presence of interstellar contributions to the EW of the C IV blend, for this feature the empirical calibration (Eq. 6.11) is expected to give better results.

Figure 6.8 shows that the strengths of suitable absorption features in the spectra of starburst galaxies is clearly related to the chemical composition. However, the equivalent widths calculated with different spectral synthesis codes for a given abundance and the same assumptions on the stellar ages tend to differ by factors up to two [347]. Therefore, at present, absolute metallicity derivations based on stellar (or stellar wind) absorption lines are expected to have errors of the same order. On the other hand, derivations of abundances differences between different galaxies observed with the same equipment and with equivalent widths measured in exactly the same way are expected to be significantly more accurate.

Passively evolving galaxies without star formation have no H II regions. Therefore, in this case there are no nebular emission lines which could be used for chemical abundance derivations. But elliptical and S0 galaxies have rich stellar absorption-line spectra, which in high-redshift objects can be observed by means of IR spectroscopy. As in the case of starburst galaxies, the equivalent widths of various diagnostic absorption features can be used to derive chemical abundances. However, as in the case of the continua, the line strengths often depend on the metallicity as well on the age of the stellar population. If the stellar ages can be constrained, the absorption spectra of the passively evolving galaxies can be used to derive accurate chemical abundances (see, e.g., [81] and [431]).

6.8 Measuring Interstellar Gas and Dust Properties

6.8.1 Properties and Kinematics of the Gas

As pointed out in Sect. 6.7, a comparison of observed emission line spectra and photoionization models provide a reliable method for deriving the chemical composition of the ionized interstellar medium of galaxies. Since H II regions result from the ionization of cool interstellar matter by short-lived massive ($M > 10M_{\odot}$) stars (and have lifetimes as short as those of the massive stars), the chemical composition of the H II gas can be assumed to be characteristic of all the interstellar gas, including the cool, neutral gas of the galaxies.

Using temperature and density sensitive emission line ratios, the nebular spectra can also be used to derive electron temperatures and densities of the ionized gas. Again using the ions of oxygen, the electron temperature can be determined from the ratio of the total flux in the two lines [O III] 4959 Å and [O III] 5009 Å and the flux of the line [O III] 4363 Å. Since the first two emission lines are due to transitions from the $^1D_2$ energy level of [O III] to the (^3P) ground state, while the line at 4363 Å results from a transition of the higher $^1S_0$ state to the $^1D_2$ state, this ratio is very sensitive to the electron temperature and independent of other parameters. Graphs

relating this (and other) line ratio(s) with the temperature can be found in [390]. The electron density in H II regions can be determined, e.g., from the ratio of the two [O II] doublet components at 3729 Å and 3726 Å, which originate from the same upper level, but have different transition probabilities, and thus are affected differently by collisional depopulation of the excited level. Corresponding graphs can again be found in [390].

Densities of the cool interstellar gas can be estimated by measuring column densities of non-saturated interstellar absorption lines and by making assumptions on the geometric distribution of the gas (e.g., assuming conditions similar to those observed in galaxies of the local universe). However, these assumptions normally cannot be verified. As a result, such estimates are not very accurate.

In principle, neutral hydrogen column densities and masses can be measured by means of the 21-cm line emission from gas-rich galaxies (For the technical details of 21-cm line observations see, e.g., [441].) If a galaxy happens to be in front of a background radio source, the 21-cm absorption can be used to measure the spin temperature of the neutral hydrogen. From the spin temperature the kinetic temperature of the neutral gas can be estimated. Unfortunately, the presently available instruments normally do not have the sensitivity to detect the 21-cm line of high-redshift galaxies. But future instruments (such as the "Square Kilometer Array," see Sect. 12.1.1) are expected to be capable of carrying out such measurements.

Temperatures and masses of the cold gas of galaxies can be derived by measuring molecular lines, such as the strong lines of carbon monoxide. The rest-frame wavelengths of the low-state CO lines are in the mm-range ($\lambda \leq 2.7$ mm, e.g., [441]). For high-redshift galaxies the CO lines are shifted to cm wavelengths where sensitive instruments with good angular resolution are available. Assuming a typical CO/H_2 ratio, such measurements can be used to estimate the total molecular gas content of high-redshift galaxies. In addition to providing mass and temperature estimates, CO line observation can provide information on the kinematics of the cool molecular gas.

CO-line emission has been successfully measured in dust-obscured high-redshift starburst galaxies, in high-redshift radio galaxies, and in high-redshift QSO host galaxies (see, e.g., [210] and the literature quoted there). Among the UV-bright high-redshift starburst galaxies, only very few have CO-line fluxes which can be detected with the sensitivities of the present-day mm-wave instruments (e.g., [28] and [112]). CO-line observations of more significant samples of high-redshift galaxies of all types will probably become feasible with the ALMA telescope which at present is under construction in Chile (see, e.g., [211]).

Since the submillimeter and radio observations mentioned above (so far) have been restricted to special and relatively small samples, most of our knowledge on the turbulent and the large-scale flows of the neutral and partially ionized interstellar gas have been derived from an analysis of the rest-frame UV absorption lines of the interstellar atoms and ions. In the low-density interstellar gas, practically all atoms and ions are in their ground state. Thus, practically all the interstellar absorption lines are resonance lines. Conveniently, the resonance lines of many common atoms and ions occur in the UV spectral range. Thus, in high-redshift galaxies these

features are observable with ground-based optical telescopes. The thermal Doppler line widths of these lines are of the order of the velocity of sound in the cool and warm gas, while the expected velocity shifts due to large-scale motions are of the order of the orbital velocities of the gas clouds in the galaxies (i.e., of the order $10^2\,\mathrm{km\,s^{-1}}$). Therefore, medium or high-resolution spectra are required for kinematic studies using these lines. With the presently available instrumentation such observations can be carried out for galaxies with redshifts up to $z \approx 4$. Examples are given in Sect. 7.1.2.

6.8.2 Deriving Dust Properties

High-redshift galaxies show a wide range of their dust content. There exist evolved as well as star forming galaxies which appear to be free of dust. In other high-redshift galaxies the UV-visual radiation is completely absorbed by a large amount of internal interstellar dust.

In Sect. 6.3 it has been explained that the presence of interstellar dust results in a characteristic modification of the UV spectra of distant star forming galaxies. Dust-free young starburst galaxies have FUV spectra with an essentially uniform continuum slope, which can be characterized by a slope parameter β (see Eq. (6.4) in Sect. 6.3). For small ages this slope changes only slightly with age. As pointed out in Sect. 6.3.2, by comparing the observed slope parameter with the value predicted for a dust-free starburst galaxy, the light attenuation and the amount of dust along the line-of-sight to the continuum emitting stars of a galaxy can be estimated.

If the properties of the dust grains are known, these data can also be used to estimate the total mass of dust along the line of sight. Since there exists little direct information on the dust properties in high-redshift galaxies, normally it is assumed that the dust has properties which are similar to those observed in the Milky Way galaxy. Because of the different conditions in the ISM of distant starburst galaxies, this assumption may not be correct.

Usually the attenuation of the light by dust is estimated using Eq. (6.5). However, as pointed out in Sect. 6.3 this relation is not always valid for high-redshift galaxies. In particular for the more heavily reddened high-redshift galaxies different laws may apply [378, 379]. Therefore, in order to obtain reliable estimates of dust masses from the reddening, the reddening law has to be determined by comparing the observed continua with those predicted on the basis of the line spectra. Some hints on the dust properties can also be obtained from observational parameters which depend on the reddening law, such as γ_{34} (see [378]).

As pointed out in Sect. 2.2.5, the UV flux absorbed by the interstellar dust is re-emitted as FIR radiation. For high-redshift galaxies this emission is redshifted into the submm and mm-wave range, where it can be observed from the ground. Thus, observation at submm/mm wavelengths provide information on the dust content of the high-redshift galaxies. As discussed in Sect. 2.2.5, the rest-frame mid and far-infrared dust emission of star-forming, dusty galaxies has a typical FIR energy distribution (Fig. 2.21). With increasing redshift the maximum of the dust emission

spectrum moves to longer wavelengths. Thus, with increasing redshift observations at millimeter wavelengths sample the spectral flux closer to the maximum of the SED. For $z < 10$ this effect approximately compensates for the decrease of the spectral flux due to the increasing distance. Therefore, up to about $z \approx 10$, the sensitivity of mm-wave observation of the dust emission is practically independent of the redshift.

Internal absorption is negligible at these wavelengths. Thus, high-redshift galaxies can be regarded as optically thin in the far infrared. Observations of the redshifted FIR emission allow us (at least in principle) to observe the total amount of dust in a galaxy, independently of the dust distribution. Using models of the dust grains (based on observations of the local ISM) the total mass of the FIR emitting dust can be estimated from the FIR luminosity (e.g., [287]). Except for very small grains, the individual dust grains radiate a continuous spectrum which, to first approximation, resembles a Planck spectrum. If the properties of the dust grains are known, the mass M_d of the radiation-emitting dust can be derived from the relation

$$M_d = \frac{F_\nu D_l^2}{(1+z)\kappa(\nu_{rf})B(\nu_{rf}, T_d)} \tag{6.12}$$

where F_ν is the observed spectral flux, D_l is the luminosity distance, $\kappa(\nu_{rf})$ is the mass absorption coefficient of the dust at the rest-frame frequency ν_{rf}, B is the Planck function, and T_d is the temperature of the dust grains (for details see, e.g., [229] and [385]). κ_ν is reasonably well known for dust in the Milky Way galaxy. At a wavelength of 0.125 mm we have about $\kappa_\nu = 19\,\mathrm{cm}^2\,\mathrm{g}^{-1}$. For estimating the dust masses in high-redshift galaxies it is usually assumed that the absorption coefficient does not depend on the redshift. This assumption may not be correct. Thus, dust mass estimates in the literature may contain systematic errors.

6.9 Probing the Distant Intergalactic Medium

A discussion of the intergalactic medium as such is outside the scope of this book. However, because of the interactions between the IGM and high-redshift galaxies, for the understanding of the physics of these objects the intergalactic environment is often important and has to be taken into account. Our main source of information about the distant IGM is the analysis of the absorption lines produced by the intergalactic plasma. For high-redshift galaxies these lines result in a significant attenuation of parts of the (rest-frame) FUV spectrum. Since high-redshift galaxies are faint targets, the intergalactic absorption of the UV light severely affects the observations of objects at very high redshifts. On the other hand, the intergalactic absorption lines can be used as a tool to derive various basic properties of the intergalactic plasma.

6.9.1 Light Attenuation

As pointed out in Sect. 2.4 in the local universe the intergalactic gas is characterized by a very low density and a high degree of ionization. Therefore (as shown, e.g., by Fig. 2.31), in the local universe the attenuation of light by intergalactic absorption lines is a very minor effect. At the cosmic epoch at which we observe the high-redshift galaxies, the intergalactic density was higher by (about) a factor $(1+z)^3$, and the fraction of the ionized hydrogen and helium was lower than now. As explained in Sect. 3.3, for the first few 10^8 years after the emission of the CMB the intergalactic gas was practically neutral. The reionization of the IGM started only after the first massive and hot stars had been formed. Therefore, intergalactic line absorption by neutral hydrogen and by helium atoms and ions increases rapidly with increasing redshift. In the spectra of high-redshift galaxies and QSOs the spectral range shortward of the Lyα line tends to be covered with numerous sharp intergalactic absorption lines, which are produced by intergalactic clouds or filaments located between us and distant objects (Fig. 6.9). These lines are most conspicuous in the intrinsically smooth spectra of QSOs and BL Lac objects, but they are present in the spectra of all high-redshift galaxies as well. Most of the observed line absorption is due to the Lyα resonance line of hydrogen at different redshifts. In the spectra of $z > 3$ objects these lines form a dense sequence of individual lines, which is visible in Fig. 6.9. Taken together these lines are called the "Lyα forest."[3] or "Lyman forest" Since the absorbing clouds have lower redshifts than the light source, the Lyman forest always starts at the wavelength of the Lyα line in the QSO or galaxy spectrum and extends toward shorter wavelengths (see Fig. 6.9). With decreasing wavelengths the Lyα lines are joined by the higher Lyman lines and, eventually, the Lyman continuum absorption from the intervening gas clouds. As shown in Figs. 9.2 and 9.3 in Sect. 9.1.2, for objects with redshifts $z \geq 6$ the continuum shortward of the Lyα line often is completely absorbed by overlapping intergalactic Lyman lines.

In Fig. 6.9, some sharp absorption features due to intergalactic matter are also found longward of the Lyα line of the QSO. These features are due to resonance absorption lines of heavier elements, which (due to winds from the galaxies) are present in part of the intergalactic H I absorbing clouds. Lyman lines and the corresponding heavy-element lines together form "Lyman absorption systems."

As noted already, the number of lines observed in a distant object increases with the redshift. According to Kim et al. [267] for $z > 1.5$ and for Lyα absorbers with (atom) column densities between $10^{13.6}$ cm^{-2} and 10^{16} cm^{-2} the number of observed absorption systems per unit redshift interval $\mathrm{d}N/\mathrm{d}z$ increases with the redshift according to

$$\frac{\mathrm{d}N}{\mathrm{d}z} = N_{z0}(1+z)^{\gamma} \approx (9.06 \pm 0.40)(1+z)^{2.19\pm0.27}. \tag{6.13}$$

[3] Strictly speaking only lines corresponding to a H I column density $< 10^{17}$ cm^{-2} are included in this definitions. For lines corresponding to higher column densities, see Sect. 9.4.

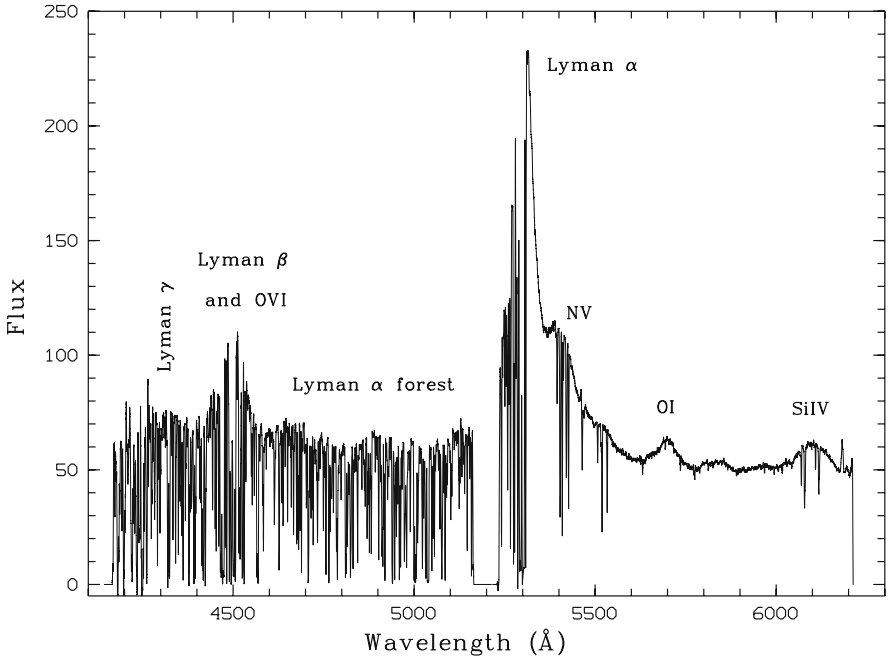

Fig. 6.9 Spectrum of the $z = 3.365$ quasar Q0103-260 showing a "Lyα forest" of sharp intergalactic absorption lines typical for the spectra of high-redshift QSOs and galaxies. The gap in the spectrum between 5164 Å and 5230 Å is an instrumental artifact. (From Frank et al. [177])

For $z < 1.5$ the differential number density $\mathrm{d}N/\mathrm{d}z$ stays almost constant.

Because of the higher cosmic density at high z, the increase of $\mathrm{d}N/\mathrm{d}z$ with z is not surprising. However, using the relations given in Sect. 3.1.5 and assuming $\Omega_M \approx 0.3$ and $\Omega_\Lambda \approx 0.7$, one finds for a non-evolving population of H I absorbers (and $z > 1.5$) an effective $\gamma < 1$. Thus, most of the increase of the number of H I absorbers with redshift is *not* caused the increasing cosmic density, but must be due to an increase of the number of H I absorbers within a comoving volume.

A related statistical quantity is the "mean H I opacity," which is defined as

$$\tau_{HI}(\lambda) = -\ln(L_\lambda / L_{\lambda 0}) \tag{6.14}$$

where L_λ is the observed spectral luminosity and $L_{\lambda 0}$ is the emitted spectral luminosity in a suitable wavelength interval around the wavelength λ in the Lyman forest. According to [267] for $z > 1.5$ this quantity increases with the redshift z approximately according to

$$\tau_{HI}(\lambda) = (0.0030 \pm 0.0008)(1+z)^{3.43\pm0.17}. \tag{6.15}$$

6.9.2 Line Profile Analysis

High-resolution spectra of the Lyman forest show that the line optical depth $\tau_{line}(\nu)$ of the non-saturated individual lines is with good approximation proportional to a Voigt profile function $V(\nu)$ given by

$$V(\nu) = \int_{-\infty}^{+\infty} L(\nu - \nu\prime)\, D(\nu\prime)\mathrm{d}\nu\prime \tag{6.16}$$

where

$$L(\nu) = \frac{\gamma}{(2\pi\,\Delta\nu)^2 + (\gamma/2)^2} \tag{6.17}$$

describes the damping contribution to the line broadening (γ being the damping constant), and where

$$D(\nu) = \pi^{-1/2}(\Delta\nu_D)^{-1} \exp -(\Delta\nu/\Delta\nu_D)^2 \tag{6.18}$$

describes the broadening due to the thermal velocity and other random velocities ($\Delta\nu_D$ being the corresponding effective Doppler width).

The Voigt profile describes spectral lines accurately if the line broadening is caused by a combination of the final lifetime of an electron energy state and by collisions (resulting in the damping term), and by the thermal (or other random) motions of the absorbing atoms or ions. For thermal motions the Doppler width is given by

$$\Delta\nu_D = \nu_c c^{-1}(2kT/m)^{1/2} \tag{6.19}$$

where ν_c is the frequency of the line center, c is the velocity of light, k is the Boltzmann constant, m is the particle mass, and T is the kinetic temperature. If an additional random velocity field (e.g., due to turbulent motions) is present, Eq. (6.19) has to be replaced by

$$\Delta\nu_D = \nu_c c^{-1}(2kT/m + v_t^2)^{1/2} = \nu_c c^{-1} b \tag{6.20}$$

where v_t describes the additional velocity field, and where $b = (2kT/m + v_t^2)^{1/2}$ is called the "broadening parameter."

If $v_t = 0$ the derivation of b from line profile fits can be used to derive directly the kinetic temperature of the absorbing gas. If $v_t \neq 0$, its value has to be determined before the temperature can be estimated. Since observed distribution of b values in Lyman-forest lines often shows a marked lower cut-off, sometimes this cut-off is used to estimate the gas temperature.

A detailed discussion of the Voigt profile fitting of Lyman forest lines can be found, e.g., in [91]. Efficient computer codes for Voigt-profile fitting are publicly available (see, e.g., http://www.ast.cam.ac.uk/rfc/vpfit.html).

Chapter 7
Observed Properties of High-Redshift Galaxies

This chapter summarizes the observational results obtained up to now for galaxies with redshifts $z > 2$. Historically, the first galaxies to be detected at such redshifts were (rest-frame) UV-bright starburst galaxies. Therefore, starburst galaxies have been investigated much more thoroughly than other types of high-redshift objects. Only recently the interest has shifted toward the passively evolving high-redshift galaxies, which became accessible due to new space-based infrared instrumentation. These high-redshifts systems are of particular astrophysical interest since (as pointed out in Sect. 2.2) in the local universe the passively evolving ellipticals and the spheroidal bulges of the disk galaxies contain most of the cosmic stellar mass. Therefore, the history and origin of these systems are among the key questions of astrophysics.

In separate sections at the end of this chapter the information available on the host galaxies of high-redshift QSOs and on the host galaxies of distant Gamma-ray bursts will be discussed. QSOs and Gamma-ray burst are rare but luminous phenomena. Thus, their host galaxies are typically observed at high-redshift. Moreover, the hosts of these objects represent high-redshift galaxies with special properties. While the host galaxies of the QSOs are among the most massive galaxies of their epochs, GRBs have been found mainly in galaxies of rather low mass.

7.1 Starburst Galaxies

Long before the first high-redshift galaxies were discovered, it had been anticipated that starburst galaxies (see Sect. 2.2.5 and Figs. 2.18 and 2.19) would in all probability be the first stellar systems to be found in the distant universe. As noted in Sect. 5.1.3, in a pioneering paper published in 1967, R. B. Partridge and P. J. E. Peebles [401] suggested that galaxies begin their lives undergoing intense bursts of star formation. Because of the presence of massive and hot stars, the newly formed galaxies were predicted to be very luminous, and to emit most of their light at far-ultraviolet wavelengths. For redshifts $z > 2$ the UV radiation is shifted to the visual or infrared spectral range, where it can be conveniently observed with ground-based optical telescopes. Based on the galaxy formation concepts of the

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_7,

1960s, Partridge and Peebles assumed that the young galaxies had masses comparable to those of present-day massive galaxies. With these assumptions they predicted the observable flux and the spectra of these objects. As the most promising way to find the high-redshift galaxies, Partridge and Peebles proposed to search for their Lyα line emission.

In 1967 sensitive astronomical detectors were limited to wavelengths $\leq 1\,\mu$m. Thus, observations of the redshifted Lyα line appeared feasible only up to redshifts where the observed wavelength of Lyα does not exceed 1 μm, which means redshifts $z = \Delta\lambda/\lambda \leq 7$. In view of this limitation, Partridge and Peebles calculated the emitted luminosity and the expected observed Lyα flux for $z \approx 7$ galaxies. Based on their calculations they predicted that the Lyα emission of very young galaxies at $z \approx 7$ should be detectable with a 90-cm telescope, equipped with an image intensifier tube (containing an S1 cathode) with exposure times as short as 5 minutes.

Prompted by the optimistic prediction of Partridge and Peebles, different groups started searches for high-redshift starburst galaxies. (For a review of these early search programs see, e.g., [203].) But although a large amount of observing time was invested, and although the surveys soon reached much fainter magnitudes than those predicted by Partridge and Peebles, for many years no normal highly redshifted starburst galaxies were found. Lyα emission was detected from distant radio galaxies and from galaxies associated with distant QSOs (e.g., [140]). But in these cases, it was not obvious whether starbursts or non-thermal ionizing sources were responsible for the observed Lyα emission. Significant numbers of high-redshift starburst galaxies were discovered only after the Lyman-break search technique (Sect. 5.1.2) had been developed.

That the early Lyα emission line searches had failed, while the Lyman-break technique was successful has two main reasons. Firstly, as discussed in Sect. 7.1.2, the combination of resonance scattering and dust absorption makes it difficult for Lyα photons to escape from gas-rich galaxies. Therefore, in many high-redshift starbursts much or all of the Lyα flux produced in H II regions is absorbed internally. As a result, on average, the Lyα emission of high-redshift starburst galaxies is much weaker than predicted by Partridge and Peebles. In fact, in some of the high-redshift starburst galaxies (the submm galaxies), dust absorption is so strong that not only the Lyα photons but practically all intrinsic UV radiation is absorbed and re-emitted at FIR wavelengths. On the other hand, most known high-redshift starburst galaxies are indeed emitters of a strong (rest-frame) UV continuum, which forms the basis of the Lyman-break technique (Sect. 5.1.2).

The second reason for the failure of the early searches was the fact that the UV-bright high-redshift starburst galaxies turned out to be much less luminous (by factors of the order 10^{-2}) than assumed by Partridge and Peebles. While Partridge and Peebles had assumed masses comparable to those of the most massive galaxies in the local universe, the observed "Lyman-break" galaxies were found to be relatively small objects of moderate mass and less extreme luminosity than predicted. As will be discussed later, these relative to the predictions lower luminosities of most (rest-frame) UV-bright starburst galaxies results from a combination of evolutionary

effects (the build-up galaxy masses from small systems by hierarchical merging) and of selection effects (as the most massive distant starburst galaxies tend to be dust-obscured).

Since the high-redshift starburst galaxies were less luminous than predicted, it also was more difficult to separate them from the numerous faint foreground objects of equal or higher apparent brightness. The Lyman-break technique provided an efficient and economic method to carry out this separation.

7.1.1 Samples of Distant Starburst Galaxies

Some of the known high-redshift starburst galaxies were found during small-scale search programs or as by-products of programs with different objectives. But the majority of the currently known high-redshift starburst galaxies were identified during a few major search projects, each of which typically yielded $\approx 10^3$ or more galaxies with $z > 2$. These projects are briefly described below. Most of the data obtained by these search projects are publicly available and form the basis for studies of the properties of the high-redshift starburst galaxies. Since the different searches used different strategies and methods, the starburst galaxies of the individual samples have different characteristics. Therefore, the different samples have to be combined to obtain an unbiased insight into the physics of the high-redshift starbursts.

The first major sample of high-redshift starburst galaxies was identified by C. C. Steidel and coworkers using the Lyman-break technique (Sect. 5.1.2). In their original survey Steidel et al. identified more than 2300 galaxies in 17 different fields (including the HDF-N) meeting the photometric criteria for $z \approx 3$ starburst galaxies. For about half of these objects spectroscopic redshifts were obtained. About 950 objects in this sample turned out to be spectroscopically confirmed star-forming galaxies in the redshift range $2 < z < 4$. The mean redshift of the spectroscopically confirmed Steidel et al. sample is $z = 2.96$. The distribution of the redshifts is roughly Gaussian around this mean, with a standard deviation in z of 0.29.A comprehensive description of this sample has been given in [498]. More recently the Steidel et al. sample has been extended significantly. It now contains about 15 000 candidates with likely redshifts $z \approx 3$ and about 2000 spectroscopic redshifts in the range $1.9 < z < 3.4$ [429].

Many more high-redshift galaxies were found during the various "deep field" surveys mentioned in Sect. 5.2.3. A sample of about 1600 star-forming galaxies in the redshift range $2 < z < 5$ with accurate photometric redshifts(m.e. $\approx 0.03(1+z)$) was derived from the FORS Deep Field (FDF) survey (e.g., [189]). Spectra were observed for about 6% of the high-redshift FDF galaxies [377].

An important sample of high-redshift starburst galaxies resulted from the Subaru Deep Field (SDF) program. Using the Lyman-break technique more than 4500 galaxies with likely redshifts $z \approx 4$ and $z \approx 5$ were identified [395, 572]. Spectroscopic redshifts were obtained for about 2% of the SDF $z > 3$ galaxy sample. These spectroscopic redshifts confirm well the photometrically predicted redshift ranges

and their distributions. Thus, although individual redshifts exist for only relatively few SDF galaxies, the SDF sample forms a valuable data base for statistical studies.

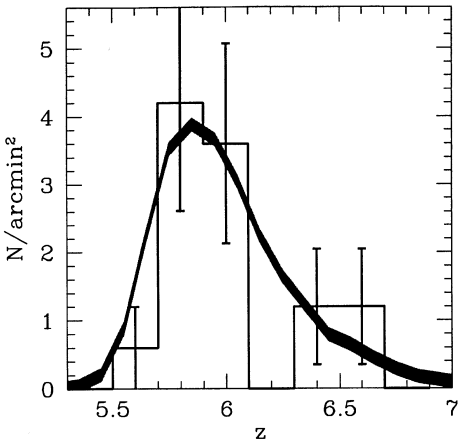

Fig. 7.1 Computed redshift distribution of the HUDF dropout sample of $z \approx 6$ starburst galaxies of Bouwens et al. ([64]) (*thick line*). For comparison the (correspondingly scaled) redshift distribution of galaxies with spectroscopic redshifts (obtained with the HST ACS grism mode by S. Malhotra et al. ([334])) for HUDF galaxies selected according to the same criteria is also included (*histogram*). (Figure from Bouwens et al. [64])

Samples of starburst galaxies at higher redshifts have been obtained during searches for i-dropout objects in the Hubble Ultra Deep Field, the HUDF Parallel Fields (see Sect. 5.2.3) and two GOODS fields observed with the HST ACS camera. The analysis of these data by several different groups resulted in a sample of >600 galaxies with redshifts $z \approx 6$ [64]. Again, individual redshifts are known for only few members of this sample. However, because of its carefully determined redshift distribution function (Fig. 7.1), this i-dropout sample provides an important data base for statistical studies of LBGs at this redshift. Examples of z-dropout galaxies from the sample of Bouwens et al. are reproduced in Fig. 7.2. Information on the individual objects in Fig. 7.2 can be found in [62].

Combining HUDF and UDF05 data yielded additional >200 galaxies with $z \approx 5$ ([382]). More than 300 galaxies with individual redshifts $2 < z < 5$ were identified during spectroscopic follow-up programs of the GOODS survey [428, 542]. In the same redshift range 186 galaxies were identified spectroscopically in the course of the VIMOS VLT Deep Survey (VVDS) [307], while about 500 galaxies with $2 < z < 3$ were identified spectroscopically during the first runs of the zCOSMOS-deep survey [314]. During the coming years this survey is expected to yield in total spectra of about 10^4 galaxies in the redshift range $1.4 < z < 3$.

A large sample of distant starburst galaxies identified from their Lyα emission has been reported by M. Ouchi and colleagues [394], who photometrically identified >800 likely Lyα emission galaxies (LAEs) with redshifts $3 \geq z \geq 5.7$ in the SXDS field. Spectroscopic confirmations were obtained for about 10% of this

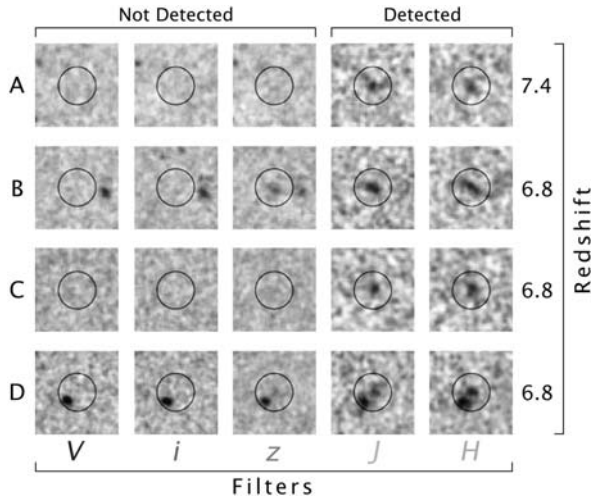

Fig. 7.2 Images of faint z-dropout (and i-dropout) galaxies observed with the ACS camera (V, i, and z-bands) and the NICMOS camera (J-band and H-band) of the Hubble Space Telescope. For each object and filter band a 3.5" × 3.5" area is reproduced. On the right the likely redshifts are given. (Credit: NASA, ESA, R. Bouwens and G. Illingworth (University of California, Santa Cruz))

sample. Another important sample of LAE candidates with $z \approx 4.5$ was derived by the "Large Area Lyα" (LALA) survey. Of these candidates, 73 could be spectroscopically confirmed as LAEs at this redshift (see [125]). Other substantial samples of high-redshift LAEs were based on the SDF (89 photometrically identified LAEs at $z \approx 5.7$ with 28 spectroscopic confirmations, see [475]), the COSMOS field (119 photometrically identified LAEs at $z \approx 5.7$ with ≈ 50 spectroscopic confirmations, see [366]), and the MUSYC survey (162 identified LAEs at $z \approx 3.1$, [195]). Most other published LAE samples suffer from relatively small object numbers, in particular at higher redshifts [248]. The largest $z > 6$ LAE sample [258] contains about 58 photometrically identified LAEs at $z \approx 6.5$ with 17 spectroscopic confirmations.

An important galaxy catalog of interest for high-redshift studies is the "GOODS Multiwavelength Southern Infrared Catalog" (GOODS-MUSIC) [208]. This catalog contains photometric data and photometric redshifts of nearly 15 000 southern galaxies of different types. It covers the redshifts range $0 < z < 5$. A part of the objects in this catalog are high-redshift starburst galaxies.

Examples of $z > 2$ dust-obscured ("submm") starburst galaxies with accurate redshifts from optical spectroscopy are listed in [94] and [246]. A sample of submm galaxies with photometric redshifts (based on their flux at FIR, millimeter, and radio wavelengths) can be found in [21].

At present many new surveys are being executed, prepared, or planned. As a result, during the next decade for all types of distant starburst galaxies the available samples are expected to grow significantly.

7.1.2 Properties of UV-Continuum Selected Starburst Galaxies

This section describes the observed properties of those high-redshift starburst galaxies which have been identified on the basis of their rest-frame UV continua. Included are the Lyman-break galaxies and most high-redshift starburst galaxies identified by means of photometric redshifts or as dropouts in the blue and visual spectral bands. Not included in this section are the high-redshift starburst galaxies identified by their Lyα emission lines. These objects are discussed separately in Sect. 7.1.3.

7.1.2.1 Appearance in Ground-Based Images

Figure 7.3 shows examples of broadband (BRI) optical images of galaxies in the redshift range $0.3 < z < 5.2$. As expected, among the objects reproduced in Fig. 7.3 the galaxies with the lowest redshifts ($0.3 < z < 1$) have the largest angular

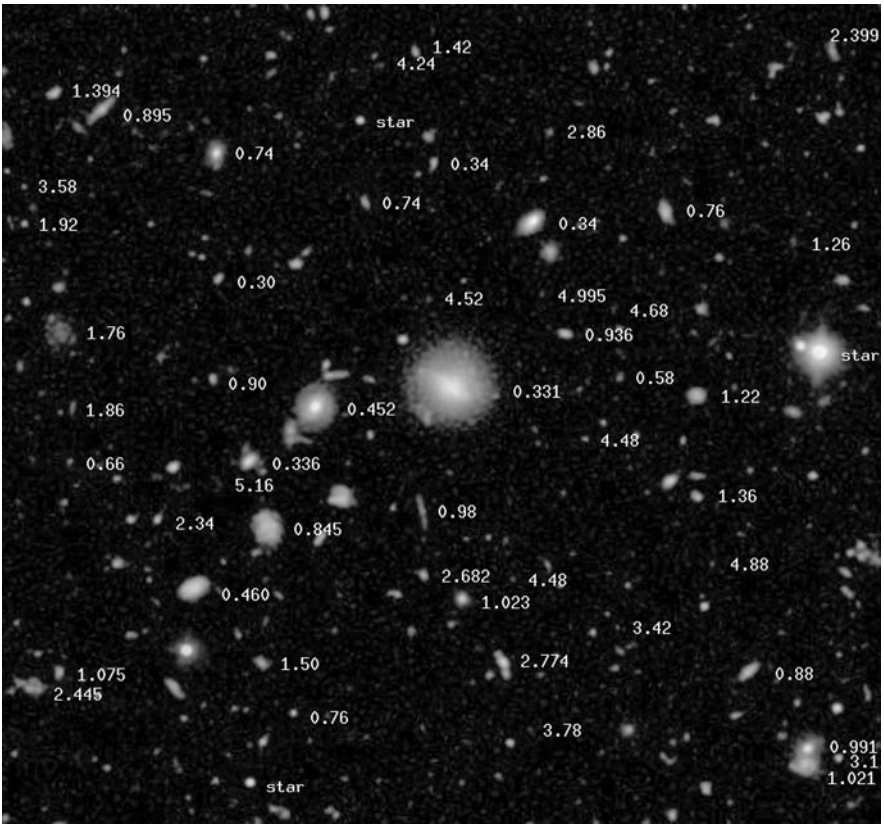

Fig. 7.3 1.5' $\times$ 1.4' section of the FORS Deep Field showing images of galaxies in the redshift range $0.3 < z < 5.2$. For part of the galaxies the redshifts are indicated by the numbers just to the right of the corresponding galaxies. Numbers with three decimals are spectroscopic redshifts, two decimals indicate photometric redshifts. (From Appenzeller et al. [17])

sizes, while the high-redshift galaxies appear small. While a decrease of the angular size in this redshift interval is to be expected, the decrease is larger than predicted for objects of constant linear size. As explained in Sect. 3.1.5 and quantified by Eq. (3.22) and Fig. 3.2, between $z = 0.3$ and $z = 2$ the angular extent of galaxies of a given linear size is expected to decrease by less than a factor of two. And the angular size is expected to increase again for higher redshifts. A simple quantitative evaluation shows that the $z < 1$ galaxies marked in Fig. 7.3 have an apparent mean angular size of about 3.5 arcsec, while 90% of the $z > 2$ galaxies have apparent sizes $<$0.8 arcsec. Obviously, the high-redshift galaxies appear smaller than expected for objects of constant linear size.

The low-redshift galaxies in Fig. 7.3 clearly represent diverse morphological types, and include ellipticals as well as disk galaxies. The morphology of the high-redshift galaxies cannot be classified reliably from ground-based images, such as Fig. 7.3. But from the observed spectra and colors it is known that all galaxies for which redshifts $z > 2$ are given in Fig. 7.3 are star-forming galaxies. The light recorded from these objects in Fig. 7.3 is redshifted UV radiation, while the low-redshift galaxies are imaged in visual light.

As noted above, most of the high-redshift galaxies have apparent angular sizes $<$1 arcsec. On the other hand, two of the high-redshift galaxies with accurate spectroscopic redshifts in Fig. 7.3 appear larger. Both are elongated objects with a maximum extent of about 3.3 arcsec (corresponding to a linear size of about 25 kpc). One of these two objects could be a large disk galaxy at $z = 2.399$. The other object (at $z = 2.445$) appears irregular. In this case we may observe a merger between two galaxies.

Better resolved ground-based images of high-redshift galaxies can be obtained using adaptive optics. Such observations have been carried out, e.g., by Akiyama and colleagues [12]. These authors used the "Infrared Camera and Spectrograph" (IRCS) of the Subaru telescope to obtain K-band images with a resolution of 0.2 arcsec (corresponding to about 1.6 kp at $z = 3$) of 31 Lyman-break galaxies with $z \sim 3$. In addition, seven Distant Red Galaxies (DRGs) were observed by this group. The observed images show the appearance of the galaxies in the rest frame visual wavelength bands. Such ground-based observations are important since with present space-based telescopes no well resolved images can be obtained longward of the H-band.

The spatial resolution of the AO-based images obtained by Akiyama and colleagues is still too low to provide detailed morphological information. However, Akiyama et al. were able to fit Sérsic profiles to the light distributions of the observed galaxies. Practically all the $z \sim 3$ galaxies (including the DRGs) showed approximately exponential (Sérsic index $n < 2$) profiles which are characteristic of local disk galaxies. About 1/3 of the galaxies in the Akiyama et al. sample show clear evidence for multiplicity or mergers.

7.1.2.2 Morphologies from HST Images

Since all ground-based surveys suffer from a very limited angular resolution of the high-redshift galaxies, most quantitative results on the morphologies and sizes

of $z > 2$ galaxies are based on deep-field observations with the Hubble Space Telescope. Rest-frame UV as well as rest-frame visual and blue HST images of high-redshift galaxies have been obtained. Examples are reproduced and described (e.g.) in [35, 64, 400], and [170]. The rest-frame-UV images were obtained with the WFPC and ACS cameras, while the rest-frame visual or blue images were taken in the H and J bands of the NICMOS NIR camera of the HST. These images have been compared with UV images and visual images of local galaxies and with simulated HST images of galaxies of various types and redshifts (as derived by [321]). The main result of the morphological studies of the HST images is a marked increase of the fraction of irregular galaxies with increasing redshift. In the local universe the great majority of the galaxies with luminosities comparable to the known high-redshift galaxies are regular disk galaxies or elliptical galaxies. Only a few percent of the bright local galaxies are irregulars. At the epoch corresponding to $z = 1$ the regular disk galaxies and elliptical galaxies still form a majority among the absolutely bright galaxies (e.g., [400]). But above $z = 2$ the fraction of regular disk galaxies and elliptical galaxies in the observed HST samples is decreasing rapidly with increasing redshift [170]. According to Beckwith et al. ([35]), in the HUDF there are "few if any galaxies at redshifts greater than ≈ 4 that resemble present-day spiral or elliptical galaxies." This is illustrated by Fig. 7.4, where examples of high-redshift galaxies observed in the HUDF are reproduced.

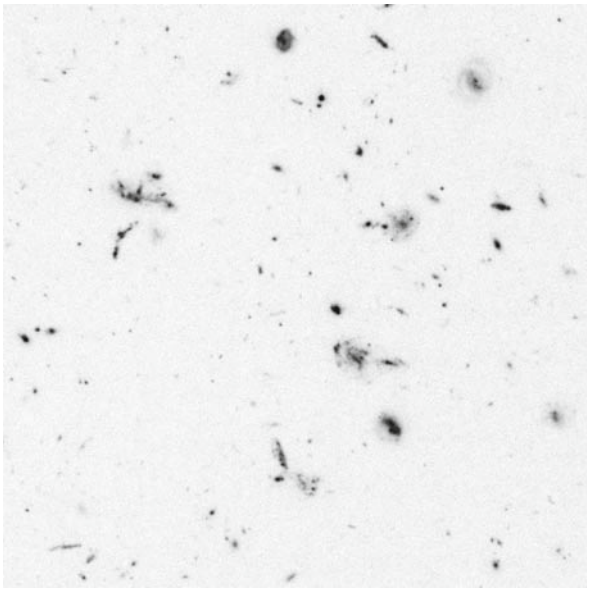

Fig. 7.4 Examples of images of high-redshift galaxies observed in the HUDF. (Credit: NASA, ESA, S. Beckwith (STScI) and the HUDF TEAM)

7.1.2.3 Galaxy Sizes from HST Observations

In connection with Fig. 7.3 it has been shown that even a cursory analysis of images of high-redshift galaxies indicates an increase of the galaxy's physical sizes with decreasing redshift or increasing cosmic time. Since the time evolution of the dark matter halos is characterized by a steady growth of their average size (see Sect. 3.3), and since the galaxy sizes are expected to scale with the sizes of their dark halos, the size increase with time is expected. Accurate measurements of galaxy sizes as a function of redshift obviously provide an important test of the theory. However, a reliable quantitative evaluation of the corresponding data is a complex task, since selection effects and technical problems can cause misleading results. One problem is the fact that at a given redshift the sizes of the individual galaxies depend on their mass and their morphology. Therefore, only galaxy samples with the same average mass (or, assuming a fixed M/L ratio, the same average luminosity) and the same morphology should be compared. Samples with comparable luminosities can be constructed, if the redshifts are known. Matching the morphological types is more problematic, since (as noted above) the morphologies of the very distant galaxies differ from those of the local universe. Finally, for $z > 5$ usually no accurate individual (spectroscopic or photometric) galaxy redshifts are available. Therefore, at such high redshifts drop-out samples have to be used, for which only the mean redshifts and the redshift distributions are known. In principle these mean parameters provide sufficient information for statistical purposes. However, uncertainties concerning the exact redshift distributions of such samples can introduce large errors in the derivation of redshift-dependent quantities. Moreover, the completeness of such samples depends on the galaxy size distribution itself. Thus, iterative procedures have to be used in order to get correct results.

The sizes and the size distribution of local galaxies have been discussed in Sect. 2.2.2. As explained in Sect. 2.2.2, in the local universe the galaxy size depends on the (stellar) galaxy mass and on the morphological type. The effective radii were found to range between about 0.1 kpc and about 15 kpc.

An extensive study of the size evolution for redshifts $z < 3$ has been carried out by Trujillo and colleagues ([527]). Their paper also includes references to and comparisons with earlier work on this topic. Trujillo et al. combined ground-based and HST data to study the sizes as a function of the luminosity and of the (stellar) mass at different redshifts. They used the Sérsic index n (see Sect. 2.2.1) to parameterize the morphology of the galaxies. As a main result Trujillo et al. found for a fixed mass and a Sérsic index typical of disk galaxies an increase of the effective radius with decreasing redshift $\propto (1+z)^{-0.40\pm0.06}$. For Sérsic indices typical for elliptical galaxies, the evolution of the galaxy sizes at a given mass was found to be proportional to $(1+z)^{-0.45\pm0.10}$. Since the M/L was found to increase with age, a somewhat stronger size evolution was found for a fixed luminosity. For the redshift range $1 < z < 3$ Trujillo et al. found for a fixed V-band luminosity a size evolution $\propto (1+z)^{-0.8\pm0.3}$.

Within the error limits similar results were found by Toft and colleagues [521], who studied a sample 41 galaxies of different types with redshifts $2 < z < 3.5$ using optical and IR data. They found for about half of their targets effective radii < 1 kpc, while the other half had radii $1 < r_e < 10$ kpc. Passively evolving galaxies were found to be systematically more compact than the star-forming galaxies. Akiyama et al. [12] found from their ground-based observations of a sample of LBGs with redshifts $z \approx 3$ a median half-light radius of about 1.8 kpc.

The size evolution at higher redshifts has been investigated by several different groups (e.g., [63, 64, 169]). All these studies were based on HST deep fields, in all cases the size evolution on rest-frame UV images was investigated, and only UV-bright starburst galaxies were included. All authors agree on a further decrease of the galaxy sizes between $z = 3$ and $z = 6$. Ferguson et al. [169] found for $4 < z < 6$ and a fixed luminosity a size decrease consistent with $\propto (1 + z)^{-1.5}$. Bouwens et al. ([64]), using a larger and deeper sample, found for $2.5 < z < 6$ and a fixed luminosity a size decrease proportional to $(1 + z)^{-1.1\pm0.3}$. Within the error limits this relation is consistent with that reported by Trujillo et al. for their lower redshift sample.

In all $z > 3$ samples rather small absolute values of the effective radii of the observed starburst galaxies were found. At $z \approx 6$ and an apparent AB z-band magnitude of $m_z \approx 27$ Bouwens et al. report a mean effective radius $r_e \approx 0.8$ kpc and a narrow distributions around this value (FWHM ≈ 0.6 kpc) [64]. According to Ferguson et al. [169] at $z = 5$ the shape of the size distribution is still compatible with being log-normal, as observed for the galaxy sizes in the local universe (Sect. 2.2.2).

7.1.2.4 Luminosities

Starting with the pioneering work of Steidel and colleagues [496] many different groups invested much time and effort in deriving the luminosity function of high-redshift star-forming galaxies. Although there is still some disagreement on the LF at very high redshifts, most recent studies (based on large and well controlled samples) give fairly consistent results. A comprehensive discussion of the observed data and their interpretation has been given by Bouwens and colleagues [65]. The paper of Bouwens et al. also contains references to the extensive earlier literature on this topic.

Of particular interest for the field of high-redshift galaxies is the luminosity function at (rest-frame) FUV wavelengths near 150 nm, where the energy output of the UV-bright starburst galaxies reaches its maximum. For $2 < z < 7$ the rest-frame FUV can be observed with sensitive CCD detectors from the ground as well as with the HST. For low redshifts and for the local universe the FUV luminosity of galaxies is known from observations with the GALEX satellite (Sect. 5.2.1).

Figure 7.5 shows (as examples) the observed luminosity functions at a (rest-frame) wavelength of about 160 nm for four different redshifts. At all four redshifts the luminosity function has a similar shape and is well approximated by a Schechter function (Eq. 2.8). However, from Fig. 7.5 it is evident that the LF evolves with redshift and cosmic time. The figure shows that the characteristic luminosity L^*, where

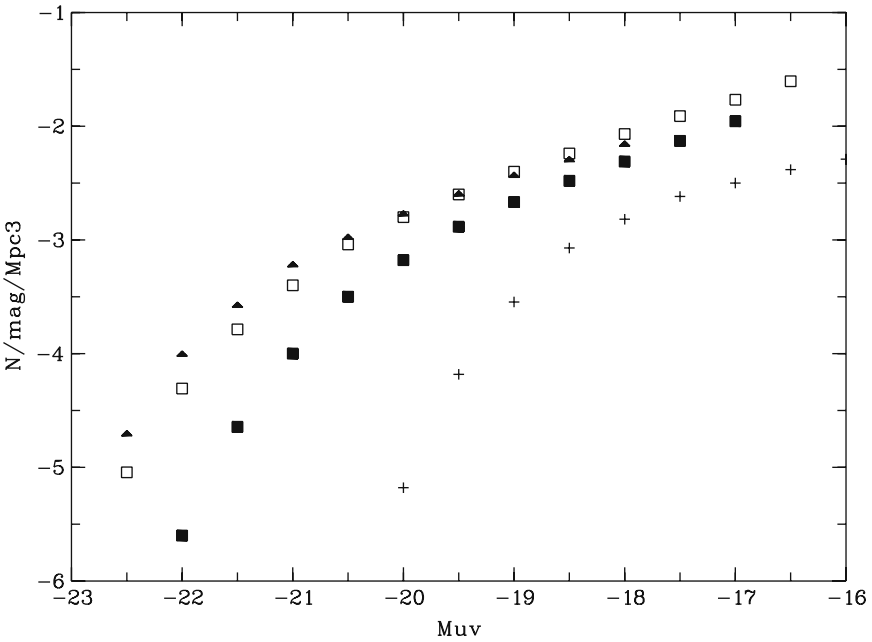

Fig. 7.5 UV ($\lambda \approx 160$ nm) luminosity functions of the UV-bright star-forming galaxies at four different redshifts. The figure shows the comoving space density (expressed as the number of galaxies per magnitude luminosity interval and per Mpc^3) as a function of their FUV luminosity, in units of absolute AB magnitudes M_{uv}. The *crosses* represent the local ($z = 0$) LF (according to [567]). The *triangles* represent the LF at $z \approx 3$ (according to [496]); the *open* and *filled squares* represent, respectively, the LFs at $z \approx 4$ and $z \approx 6$ according to [65]

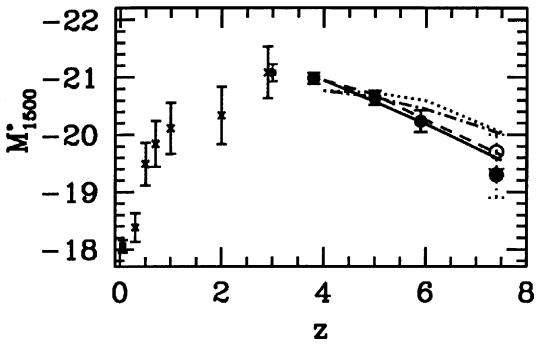

Fig. 7.6 The Schechter function parameter L^* (characteristic luminosity), expressed as an absolute magnitude M^* at 1500 Å, for samples of star-forming galaxies at different redshifts. The *dotted*, *dashed*, and *dash-dotted lines* represent different theoretical model predictions of the evolution of M^* at high redshift. The *solid line* indicates the evolution expected from the predicted growth of the halo masses with time, assuming a constant halo mass to galaxy light ratio. (From Bouwens et al. [65], where details on the derivation of M^* and references to the individual data points and models are listed)

the exponential factor of the Schechter function becomes important, is significantly larger at $z = 3$ and $z = 4$ than in the local universe. At $z = 6$ the characteristic luminosity L^* is smaller again. Quantitatively this behavior is illustrated by Fig. 7.6, where the Schechter function parameter L^* (expressed as a characteristic absolute AB magnitude M^* at the corresponding wavelength) is plotted as a function of the redshift for $0 < z < 7.5$. A more recent study, based on nearly 300 000 galaxies in the COSMOS field with photometric data in 8 filter bands, confirms the increase of M^*_{1500} between $z = 0$ and $z \approx 4$ with improved accuracy [190]. Obviously, at the epoch corresponding to a redshift $z \approx 3$ the bright galaxies were (at 1500 Å) about 15 times brighter than today. Because (due to merging and mass accretion) the galaxies have (on average) become more massive since that epoch, the higher UV luminosities indicate a much higher specific star-formation rate at $z = 3$.

The Schechter function parameters Φ^* (Normalization) and α (the faint-end slope) show less clear evolution with the redshift. According to Fig. 7.7, relative to the local ($z = 0$) data, at high redshifts both these parameters have smaller values. However, for $z > 0.5$ no statistically significant variation can be found from the data of Fig. 7.7. For $z \geq 0.4$ the data points in this figure are consistent with a constant faint-end slope $\alpha \approx -1.7$. This rather steep slope indicates that faint galaxies contribute significantly to the total UV radiation field at high redshift. More recent results on Φ^*_{1500} from the COSMOS and FDF fields ([190]) are, within the error limits, compatible with Fig. 7.7, but indicate a steady decrease of Φ^*_{1500} between $z = 1$ and $z = 4.5$.

Since the number of reliably identified galaxies with redshifts $z > 7$ is still small, the luminosity function beyond $z = 7.5$ remains uncertain at present. D. P. Stark and colleagues [494] recently identified several emission line objects which possibly are Lyα emitters with redshifts $z \approx 10$. According to Stark et al. two of their candidates are probably starburst galaxies at $z \approx 10$. Since a very small area was sampled by Stark et al., this result indicates a high abundance of bright starburst galaxies at this redshift. However, other sensitive searches for starburst galaxies at redshifts ≥ 9 (e.g. [66]) were unable to find any galaxies at this redshift. Upper limits derived from these non-detections are consistent with a further decrease of L^* or Φ^* beyond $z = 8$.

7.1.2.5 Star-Formation Rates (SFR)

Because of the relation between the UV luminosity and the star-formation rate of galaxies (Sect. 6.6), the observed luminosity functions of the high-redshift galaxies can be used to estimate the star formation rates of the individual galaxies and the cosmic star-formation rate density as a function of the redshift.

Since the SFR is related to the *intrinsic* UV luminosity, the observed luminosities have to be corrected for internal dust extinction. As pointed out in Sect. 6.3 such corrections are not very reliable at present. As a result, the derivation of the star-formation rates still involves uncertainties. Without correction for dust extinction the star-formation rates of UV-bright starburst galaxies with redshifts $z \approx 3$ are usually found to range between 10 and $10^2\, M_\odot\ \mathrm{yr}^{-1}$. If plausible corrections for

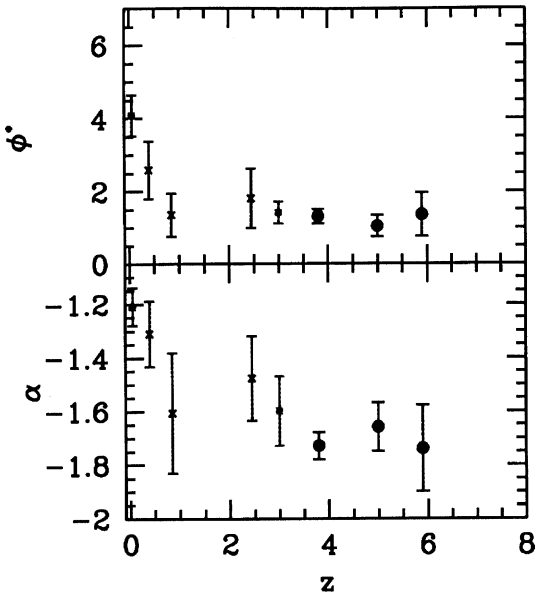

Fig. 7.7 Observed evolution of the Schechter function parameters Φ^* (normalization) and α (faint-end slope) of the UV-LF of starburst galaxies as a function of the redshift. Φ^* is given in units of 10^{-3} Mpc^{-3} comoving volume. (From Bouwens et al. [65], where references to the individual data points can be found)

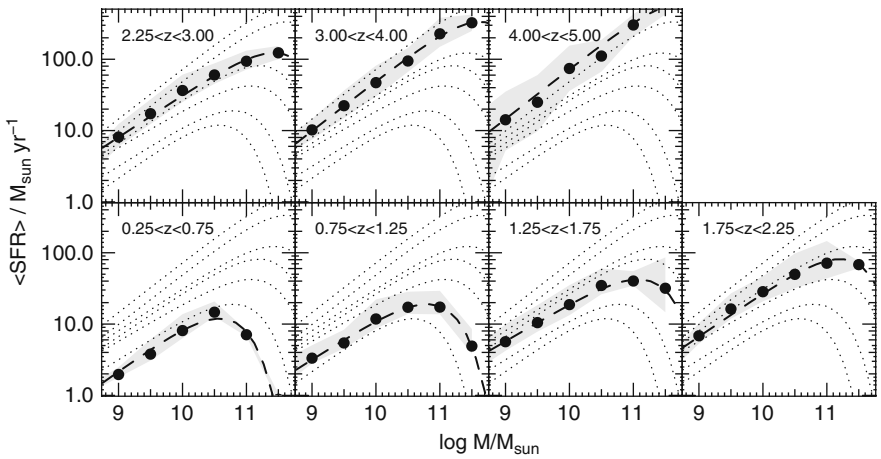

Fig. 7.8 Average star-formation rates as a function of stellar galaxy mass for different redshift intervals, as observed in the FORS Deep Field. The *shading* indicates the confidence region. (From Drory and Alvarez [146])

dust extinction are applied, in some cases SFRs up to about $10^3\ M_\odot\ \mathrm{yr}^{-1}$ are found. But more typical are dust-corrected values of the order $10^2\ M_\odot\ \mathrm{yr}^{-1}$ [203].

Figure 7.8 gives SFRs derived by Drory and Alvarez for the galaxies observed in the FORS Deep Field (FDF) as a function of the stellar mass and for different redshift intervals. As expected, for small mass values the SFR increase with increasing mass. However, at high mass values the observed SFRs start to drop again. Moreover, the maximum of the plots showing the SFR as a function of the mass systematically depends on the redshift (and cosmic time). At high redshifts the maximum occurs at high mass values, while in the local universe and at low redshifts the most massive galaxies have, on average, only small star-formation rates. In the astronomical literature this transition of the maximal star formation activity from high-mass galaxies at high redshifts to low-mass systems at low redshifts is referred to as "downsizing." It can be explained by an earlier beginning of star formation and an earlier exhaustion (or removal) of the available gas in the more massive systems.

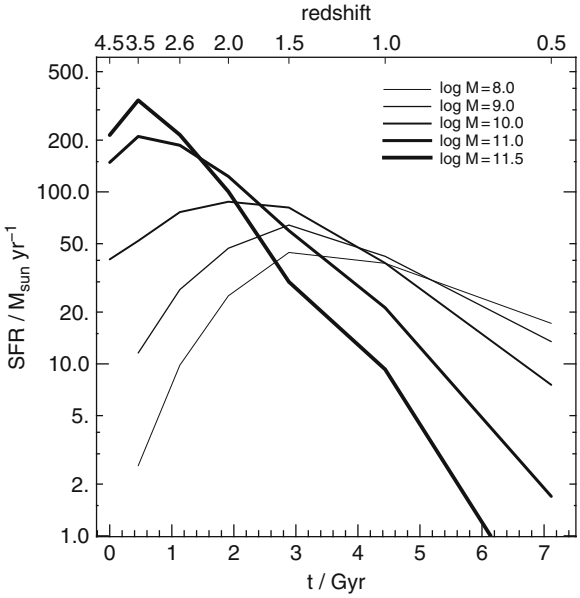

Fig. 7.9 Average star formation histories of FDF galaxies with different stellar mass M (in units of $M_\odot$) as a function of redshift and cosmic time. This plot has been obtained by integrating the SFRs given in Fig. 7.8. Note that the zero point of the time axis is set at the epoch corresponding to $z = 4.5$ (or a cosmic age of about 1.3 Gyr). (From Drory and Alvarez [146])

Another presentation of the star formation history of the high-redshift galaxies is given in Fig. 7.9. This figure, which is based on the data of Fig. 7.8, shows the average total SFRs of the FDF sample as a function of time (and redshift) for different stellar mass values. In this figure the downsizing effect is even more conspicuous. Moreover, the figure shows that for $z \leq 1.5$ the mean star formation activity decreases with time for all (investigated) galaxy masses. But the slope of the decrease is less steep for the low-mass systems.

7.1.2.6 The Cosmic Star-Formation Rate Density

If the star-formation rates of complete samples of galaxies are known, the total space density of the star formation can be derived by adding up all galaxies in a given volume. Since with increasing distance and redshift, galaxy surveys become progressively incomplete at the low-luminosity end, in practice this always involves an extrapolation or corrections at low galaxy luminosities and masses. Such corrections are possible and reliable, if the luminosity functions at different redshifts are known.

Plots showing the star-formation rate density as a function of the redshift were first published by P. Madau and colleagues [327]. Therefore, such plots are called "Madau diagrams." Madau diagrams, based on currently available large data sets have been published, e.g., in [65, 79, 429, 524]. Figure 7.10 shows Madau diagrams derived by Bouwens and colleagues [65], based on the same data sets as Figs. 7.6 and 7.7. For the conversion of UV luminosities to SFRs a constant ($> 10^8$ yr) star formation model and a Salpeter IMF were assumed. Because of the uncertainties introduced by the corrections for dust extinction, the SFRs are given with a best-estimated dust correction and without any dust correction. The uncorrected relation

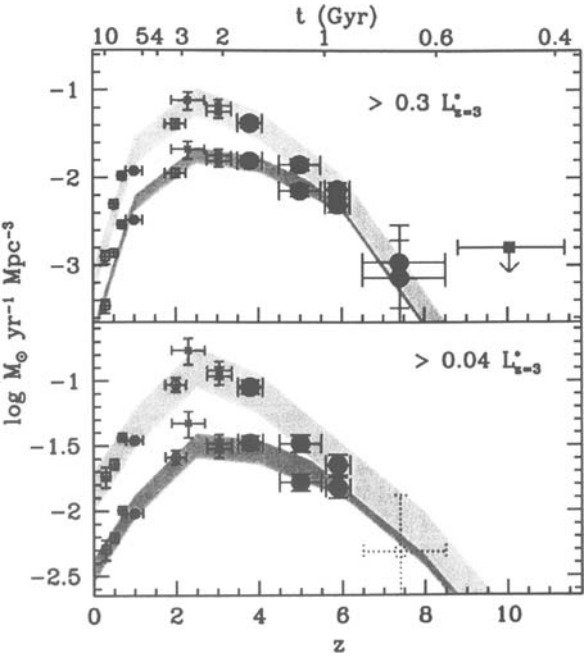

Fig. 7.10 Evolution of the cosmic star-formation rate density with the redshift (and cosmic age) according to Bouwens et al. [65]. The *upper panel* shows the SFR density if only galaxies brighter than 0.3 L^* at $z = 3$ are included. The *lower panel* gives the result for all galaxies brighter than 0.04 L^* at $z = 3$. For both cases the rates derived without dust correction (lower data points) and with dust correction (upper data points) are given. The *shaded regions* indicate systematic uncertainties. For a detailed discussion and references to the individual data points see [65]. (From Bouwens et al. [65])

can be regarded as a reliable lower limit of the star-formation rate density. Because of the uncertainties of the faint ends of the luminosity functions, the diagrams are given for two different luminosity cut-offs. The dust-corrected data in Fig. 7.10 include UV-bright as well as strongly reddened starburst galaxies. Since dusty starburst galaxies are most abundant at redshifts near $z \approx 2.5$ (see Sect. 7.1.4), the difference between the corrected and uncorrected cosmic SFRs reach a maximum near that redshift.

As shown by the figure, the cosmic star formation reached a maximum in the past at the epoch corresponding to $z \approx 3$. After the epoch corresponding to $z = 2$ the cosmic SFR decreased rapidly to the relatively low present-day value.

7.1.2.7 Mass Estimates

If the stellar populations and star formation histories can be estimated from spectra or from photometric data, stellar masses can be calculated from the observed galaxy luminosities. If redshift information is available, the dependence of the galaxy masses on the redshift can be determined and compared to the theoretical predictions. Figure 7.11 presents the observational results of such a study. The figure is reproduced from a paper by F. Elsner and colleagues [156], who determined mass values for nearly 15 000 galaxies with redshifts $0 < z < 5$. This paper also contains a comparison with—and references to—the earlier literature on this subject. The mass values were determined by fitting multi-band photometric data to a library of synthetic spectra. The photometric data were taken from the GOODS-MUSIC multicolor catalog [208], which includes photometry in 14 bands, extending from the UV to 8 μm. The catalog also lists photometric redshifts, as well as spectroscopic redshifts for a subsample of the catalog objects [208].

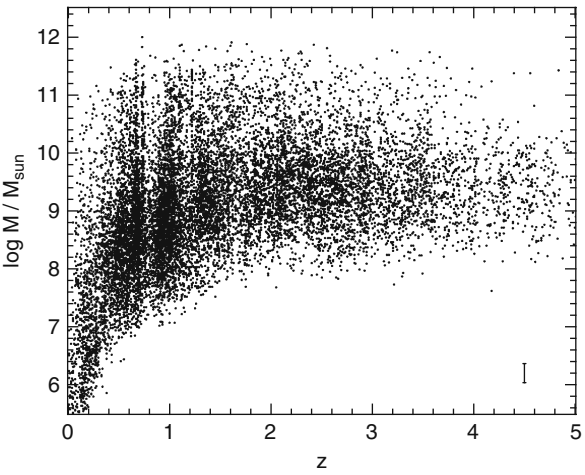

Fig. 7.11 The stellar mass of individual galaxies in the GOODS-MUSIC catalog as a function of the galaxy redshift. The bar in the lower *right-hand corner* indicates the typical error of the mass estimates. (From Elsner et al. [156])

Because of selection effects, the interpretation of Fig. 7.11 requires some caution. The lower envelope of the data points in Fig. 7.11 is entirely due to the magnitude limit of the observations. Intrinsically faint galaxies are included only at the lowest redshifts. On the other hand, the upper envelope essentially represents the upper mass limit (or, more accurately, the exponential drop-off of the corresponding Schechter mass functions) of the galaxies. Near $z = 0$ the sampled volume becomes small, resulting in statistical uncertainties at high luminosities. At low redshifts, all types of galaxies are included. At high redshift, the sample consists essentially of starburst galaxies.

Even a cursory look at the upper envelope of the distribution in Fig. 7.11 shows that the upper mass limit is increasing with decreasing redshift (or increasing cosmic time), as expected from the hierarchical galaxy evolution scenario.

Up to a redshift of about $z = 3.5$, Elsner et al. were able to fit Schechter functions to the observed mass distributions and to derive Schechter parameters (as defined Sect. 2.2.2) as a function of the redshift. According to these authors the most significant redshift evolution is found for the parameter Φ^* (normalization), which decreases between $z = 0.5$ and $z = 3.5$ by more than a factor 10. No significant variation could be detected for the faint-end slope α, while the characteristic mass (as expected) was found to decrease between $z = 1$ and $z = 3.5$

With the known mass functions (approximated by the Schechter functions) the observed part of the mass distribution can be extrapolated to low galaxy masses. In this way, the mass distributions can be corrected for the incompleteness at low mass and high redshifts. By integrating over the distribution at a given redshift (taking into account the observed volume as a function of redshift) the relation between the stellar mass density and the redshift can be calculated. Results obtained by different authors are collected in Fig. 7.12. As shown by the figure, within the error limits the different derivations agree reasonably well. All studies indicate a steady increase of the comoving stellar mass density with decreasing redshift. Obviously, this is to be expected in view of the ongoing star formation in the universe. According to Fig. 7.12 almost half of the present cosmic stellar mass had already been formed at the epoch corresponding to $z = 1$. On the other hand, at $z = 3.5$ only about 6% of the present stellar mass had already been present. At a redshift of $z = 5$ the mass accumulated in stars had reached only about 1% of the present-day value.

7.1.2.8 The Spectra of High-Redshift Starburst Galaxies

The most important source of information on the physics of distant galaxies are spectroscopic observations. With the presently available telescopes and modern spectrographs, the continuum and the absorption-line spectra of luminous galaxies can be observed routinely for redshifts up to about $z = 5$. Emission lines (allowing us to derive accurate individual redshifts) and the overall continuum energy distribution can be studied at significantly higher redshifts. But absorption line spectra of acceptable quality of $z > 5$ galaxies are restricted to exceptionally bright or gravitationally amplified objects (e.g., [143]). Apart from the low photon flux of $z > 5$ galaxies, spectroscopy of such objects is hampered by the fact that at

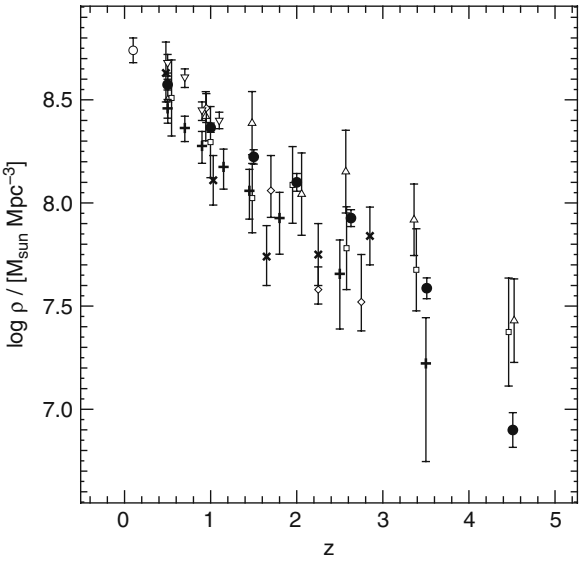

Fig. 7.12 Comoving stellar mass density as a function of the redshift. Sources of the different data points are [156] (*filled circles*), [148] (*open triangles* and *open squares*), [133] (*open diamonds*), [147] (*reversed triangles*), [174] (*tilted crosses*), [175] (*upright crosses*), [104], (*open circles*). (From Elsner et al. [156])

$z > 5$ most of the flux is shifted to the IR spectral range, where ground-based observations are restricted to a few atmospheric transmission windows, and where space-based spectroscopy (so far) has to make use of telescopes with relatively small apertures.

For the reasons outlined above, a large majority of the published spectra of high-redshift starburst galaxies have redshifts $z < 5$, and most of the available detailed spectral information is based on objects with redshifts $2 < z < 4$. From UV-bright starburst galaxies at this redshift the photon flux is still high enough to obtain low-resolution spectra for a significant range of intrinsic luminosities. Medium resolution spectra can be obtained at least for the brighter objects in this redshift range. References to samples of galaxy spectra at this redshift are given in Sects. 5.2 and 7.1.1.

Detailed studies of the properties of the rest-frame UV spectra of distant starburst galaxies have been carried out by A. Shapley and colleagues [471] and by S. Noll et al. [377]. The sample discussed by Shapley et. al. was selected using the Lyman-break technique. The sample of Noll et al. was based on photometric redshifts. The Noll et al. sample is smaller, but it covers a larger redshift range. Apart from differences resulting from the different selection criteria and the different redshift range the results of the two samples are in good agreement.

In order to represent typical spectral properties and in order to improve the S/N, often mean or "composite" spectra are calculated by averaging (or stacking) the observed spectra of a given sample. Figure 7.13 shows, as an example, the averaged

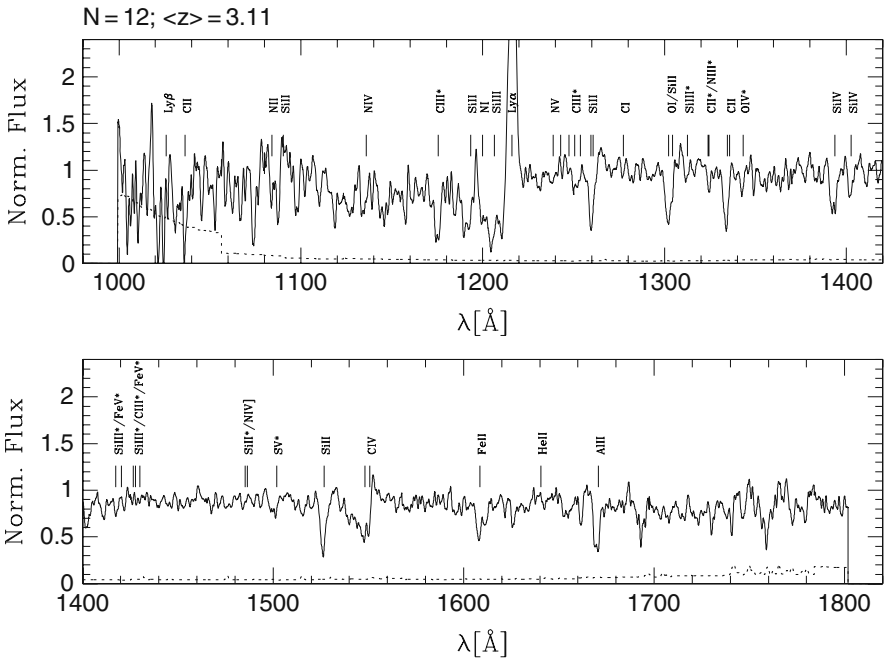

Fig. 7.13 Medium-resolution composite spectrum of $z \approx 3$ starburst galaxies. The spectrum is an average of 12 galaxy spectra with redshifts $2.4 < z < 3.4$ ($\bar{z} = 3.1$). The spectral features originating in stellar atmospheres are indicated by an *asterisk* behind the identification. The broad blueshifted absorption features are stellar-wind lines. The narrow, deep features are interstellar absorption lines. The *broken line* indicates the noise level. (From Mehlert et al. [347])

medium resolution (R = 2000) rest-frame UV spectrum of 12 starburst galaxies with a mean redshift of $z = 3.1$.

Spectra covering the rest-frame visual range of high-redshift starburst galaxies have been published by Erb et al. [159], Kriek et al. [281], Law et al. [302], Maiolino et al. [330], and others. An example is presented in Fig. 7.14. A comparison with Fig. 2.18 shows that the optical emission line spectra are similar to those observed in local starburst galaxies. A comparison of the visual range absorption line spectra is more difficult, since the continuum S/N of the ground-based NIR spectra of these relatively faint objects is generally too low to provide information on the absorption lines.

Figure 7.13 shows that the main features of the UV line spectra of starburst galaxies at redshifts $z \approx 3$ are (relatively narrow) interstellar absorption lines (of, e.g., C I, C II, N I, N II, Si II, Fe II, Al II, and Al III), broad absorption features produced mainly by the stellar winds of hot stars (e.g., Lyα, N V, C IV, Si IV, see Sect. 2.1.2 and Fig. 2.2), and relatively weak absorption lines and blends typical of the photospheric spectra of hot stars. In Fig. 7.13 the lines and blends originating in stellar atmospheres are indicated by an asterisk behind the identification. In addition to the absorption features, emission lines originating in the interstellar medium of

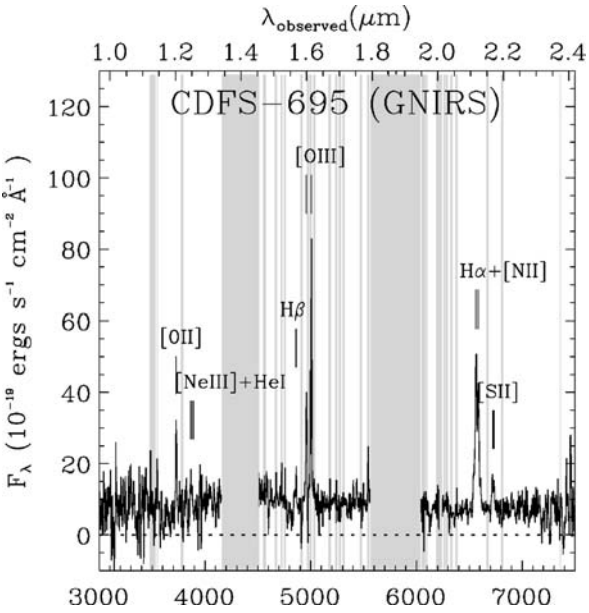

Fig. 7.14 Rest-frame visual-range spectrum of a $z = 2.225$ starburst galaxy. The gaps in the spectrum are due to the atmospheric NIR H_2O absorption. The wavelength scale at the top gives the observed wavelength (in μm). The scale at the bottom gives the rest-frame wavelength in Å. (From Kriek et al. [281])

the galaxies (e.g., Lyα, and the intercombination line C III]) or in stellar winds (e.g., C IV and He II) are usually present.

A comparison of Fig. 7.13 with Fig. 2.19 (Sect. 2.2.6) shows that the UV spectra of the high-redshift starburst galaxies resemble closely the spectra of local star-forming galaxies. Moreover, as demonstrated by many different authors, the continuum energy distribution and most of the absorption features of the spectra of individual high-redshift galaxy spectra can be reproduced well by spectral synthesis models which assume young stellar populations with initial mass functions, dust reddening laws, and other parameters as observed in the local universe (e.g., [416, 439, 470, 513]). The starburst ages are found to be of the order $\leq 10^8$ yr, and in some cases $< 10^7$ yr. These starburst ages are small compared to the age of the universe at the observed cosmic epochs.

As pointed out in Sect. 3.3, there are reasons to expect that certain parameters, such as the initial mass function and the dust properties, may have been different at the epochs at which the light of the high-redshift galaxies was emitted. In principle, these quantities can be included as free parameters in the model fits. However, with too many free parameters the fits to population synthesis models become impossible to interpret. Thus, although the observations of the high-redshift galaxies are consistent with stellar population properties similar to those observed in local starburst galaxies, the exact parameters of the stellar populations are not well constrained.

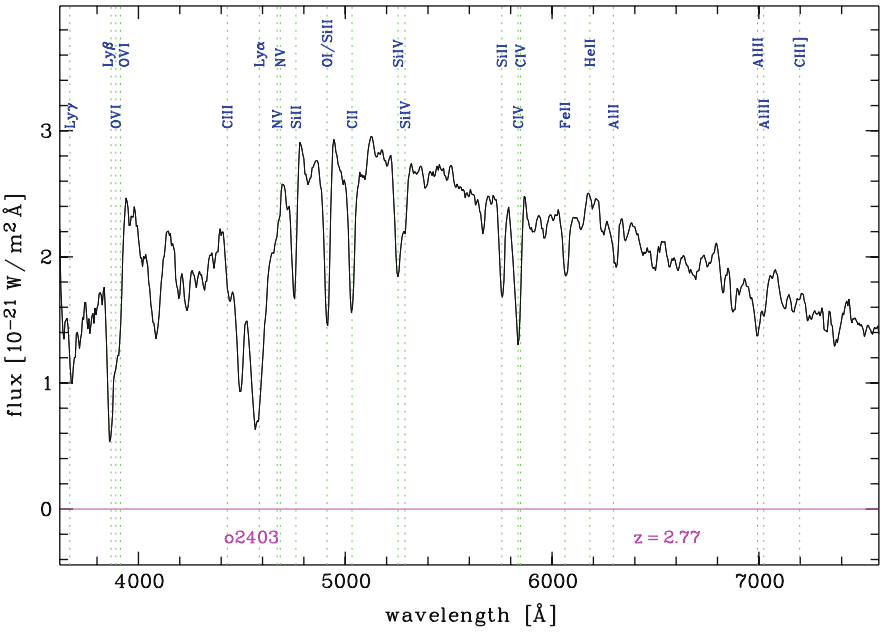

Fig. 7.15 Low resolution spectrum of the starburst galaxy FDF-5903 ($z = 2.774$) showing the presence of a significant He II 1641 Å emission line. The line (redshifted to 6193 Å) has a rest-frame emission equivalent width of about 1.5 Å. (Spectrum courtesy S. Noll)

Part of the spectra of the $z \approx 3$ starburst galaxies show significant He II 1641 Å emission . Although relatively weak in the spectrum of Fig. 7.13, the line is clearly visible in the spectrum reproduced in Fig. 7.15. The line is also present in the composite spectrum reproduced in Fig. 5.1 (Sect. 5.1.2). At least in the majority of the high-redshift starburst galaxies, the observed strength of this line can be explained by the presence of Wolf-Rayet stars. The strength of the 1641 Å feature varies strongly between individual high-redshift galaxies. Because of the short lifetime of WR stars, this variation probably reflects the corresponding scatter of the starburst ages of the observed high-redshift galaxies. However, the varying He II strength could also be caused by variations of the IMF or the chemical abundance (see, e.g., [452]).

While the absorption spectrum in Fig. 7.13 is fairly typical for starburst galaxies at this redshift, due to the selection criteria of the particular sample, the Lyα emission is atypically strong. A more representative Lyα emission strength can be seen in Fig. 5.1. Between the individual high-redshift galaxies Lyα profiles show large variations. The observed profiles range from almost pure absorption (e.g., Fig. 7.26) to very large emission equivalent widths (e.g., Fig. 7.35). Figure 7.16 shows the observed Lyα equivalent width distribution for a typical sample of continuum-selected starburst galaxies at $z \approx 3$. As demonstrated by this figure, most UV-bright starburst galaxies of this redshift show rather modest (emission or absorption) Lyα

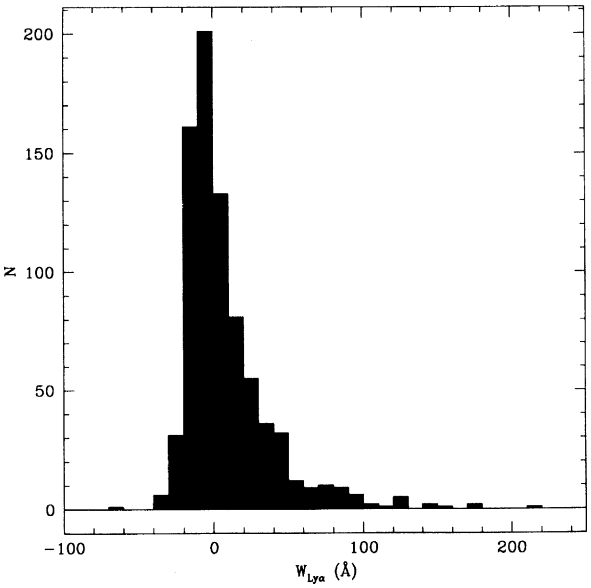

Fig. 7.16 Distribution of the Lyα emission equivalent widths of Lyman-break galaxies with redshifts of $z \approx 3$. (From Shapley et al. [471])

equivalent widths. The median value of the distribution is very close to $W_{Ly\alpha} = 0$ (corresponding to a line profile where the emission and the absorption components compensate each other). Only a small minority of the known galaxies have Lyα emission equivalent widths > 20 Å. The properties of these strong Lyα emitters (referred to as "LAEs") differ from those of other UV-bright starburst galaxies. Therefore, the LAEs are discussed separately in Sect. 7.1.3.

At first glance the large variation of the profiles of the Lyα line appears surprising. As illustrated by Figs. 7.14 and 2.18, the optical-range spectra of starburst galaxies always include strong emission lines of the Balmer series of hydrogen. These Balmer emission lines result from the recombination of hydrogen which is ionized by the UV radiation of the hot young stars of the starburst galaxies. During the recombination much of the original stellar UV radiation is converted into Lyα photons. Thus, internally an intense Lyα radiation field must be present in all starburst galaxies.

That this intrinsic Lyα radiation field normally is not observed is due to the high absorption cross-section of hydrogen atoms for Lyα radiation. Even a small amount of neutral hydrogen can make a galaxy completely opaque for Lyα photons. Normally, the absorption of Lyα by an interstellar hydrogen atom will be followed by a re-emission at the same frequency. Thus, hydrogen atoms essentially scatter the Lyα photons, which, in principle, still can escape from the surface of a gas cloud after a random walk through the neutral hydrogen containing layer. However, the resonance scattering greatly increases the effective light path of these photons. As a result, even a small amount of absorbing material embedded in the

scattering layer results in the eventual capture of the Lyα photons. The most efficient UV absorber is interstellar dust, which is normally abundant in star-forming regions.

The combination of resonance scattering by neutral hydrogen and dust absorption obviously reduces the escape probability of Lyα photons and provides a plausible explanation for the absence of Lyα emission in the spectra of many starburst galaxies. Qualitatively, this is confirmed by model computations of the radiative transfer of the UV radiation in these galaxies. As shown, e.g., in [453] the observed absorption profiles of the high-redshift starburst galaxies can be reproduced well by models of the type outlined above.

Figure 7.17 shows composite spectra of $z \approx 3$ starburst galaxies with different Lyα properties, ranging from almost pure Lyα absorption to emission equivalent widths > 20 Å. As shown by the figure, the LAEs (emission $EW > 20$ Å) which are represented by uppermost composite spectrum in Fig. 7.17, show peculiarities. But the mean spectra of the other subsamples are qualitatively similar and show only subtle differences.

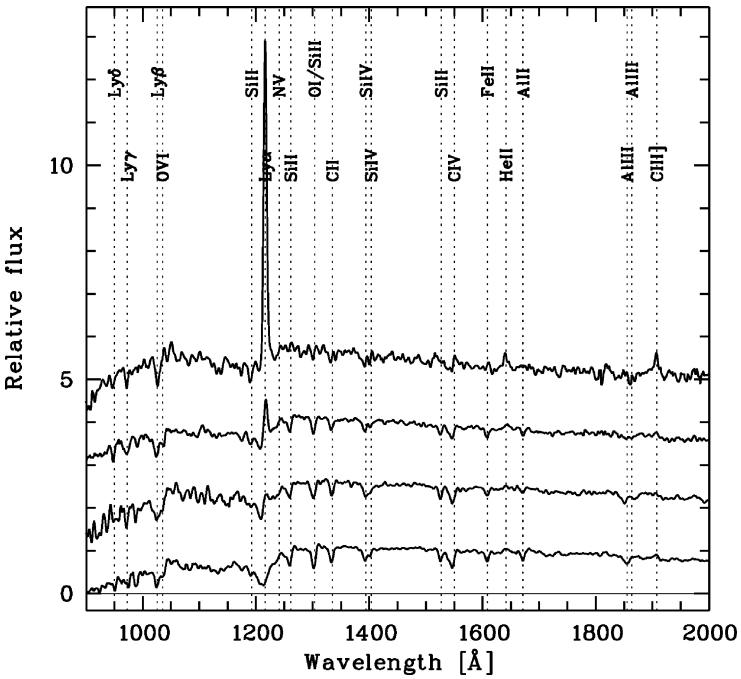

Fig. 7.17 Low-resolution composite spectra of $z \approx 3$ samples of starburst galaxies with different Lyα emission strength. The uppermost composite spectrum represents galaxies with rest-frame Lyα equivalent widths > 20 Å (i.e., "LAEs," see Sect. 7.1.3). The three lower spectra represent galaxies with, respectively, $2 < EW_{Ly\alpha} < 20$ Å, $EW_{Ly\alpha} < 2$ Å, and no detectable Lyα emission. The spectra are normalized to have equal flux at 1500 Å, and the upper three spectra are shifted by multiples of one flux unit. (From Noll et al. [377])

An accurate quantitative analysis reveals that there exist various correlations between the Lyα emission strength and other spectroscopic properties. Among the most conspicuous correlations is a marked decrease of the strength of low-ionization interstellar absorption lines with increasing Lyα emission equivalent width. This is demonstrated by Fig. 7.18, where the average equivalent widths of six strong low-ionization interstellar (LIS) absorption lines are plotted as a function of the Lyα equivalent width. The interstellar absorption lines included in the average are Si II 1260 Å, OI/SiII 1303 Å, C II 1334 Å, Si II 1526 Å, Fe II 1608 Å, and Al II 1670 Å. Since these lines are relatively strong and not significantly blended, they are generally used to represent the overall behavior of the LIS lines. Figure 7.18 also shows that the observed anticorrelation of the LIS strength and Lyα emission does not depend on the redshift. Other quantities which correlate with the Lyα emission are the equivalent width of the C IV stellar-wind absorption features, the UV continuum slope, and the He II 1641 Å, O III 1663 Å, and C III 1909 Å emission lines [377, 471]. Moreover, the Lyα emission equivalent widths is found to decrease with increasing UV continuum luminosity.

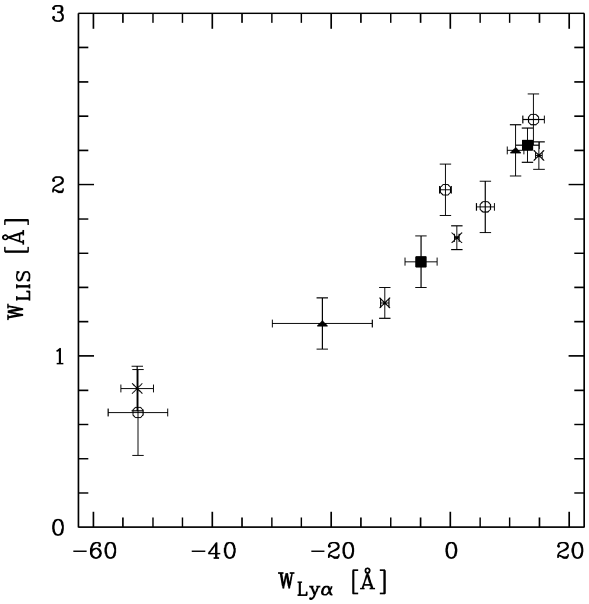

Fig. 7.18 The strength of the low-ionization interstellar absorption lines as a function of the Lyα equivalent width in the spectra of $2 < z < 4$ starburst galaxies. Negative values of the equivalent width indicate line emission. W_{LIS} represents the average equivalent width of the six strong low-ionization interstellar (LIS) lines Si II 1260 Å, OI/SiII 1303 Å, C II 1334 Å, Si II 1526 Å, Fe II 1608 Å, and Al II 1670 Å. The *crosses* are based on the data of Shapley et al. ([471]), the *open circles* indicate the corresponding values derived by Noll et al. ([377]). The *filled squares* and *triangles* represent subsamples with $2 < z < 3$ and $3 < z < 4$, respectively. (From Noll et al. [377])

According to Noll et al. [377] in the redshift range $2 < z < 4$ the average Lyα emission equivalent width increases with the redshift. In the sample used by these authors the galaxies with $2 < z < 3$ show, on average, Lyα absorption, while the galaxies with $3 < z < 4$ show an average Lyα emission. Moreover, Noll et al. found the fraction of galaxies with significant ($EW \geq 2$ Å) Lyα emission to increase from 26% at $2 < z < 3$ to 67% at $3 < z < 4$ and to 73% at $4 < z < 5$. An increase of the Lyα emission with increasing redshift is also indicated by the results of Vanzella and colleagues [541] who found significant Lyα emission in the spectra of 13 galaxies (72%) of a sample of 18 continuum-selected galaxies with $z > 5$.

Since the samples of Noll et al. and of Vanzella et al. contain normal starburst galaxies as well as LAEs, their results could in part be due to a higher frequency of LAEs at high redshifts. This cannot be excluded since, as explained in Sect. 7.1.3, the redshift evolution of the relative frequency of LAEs is not reliably known at present.

Figure 7.19 illustrates the redshift evolution of the overall spectral properties of the UV-bright starburst galaxies in the redshift range $2 < z < 4$. Apart from the marked increase of the Lyα emission, the higher-redshift spectra show systematically weaker interstellar lines and weaker stellar-wind absorption features. More-

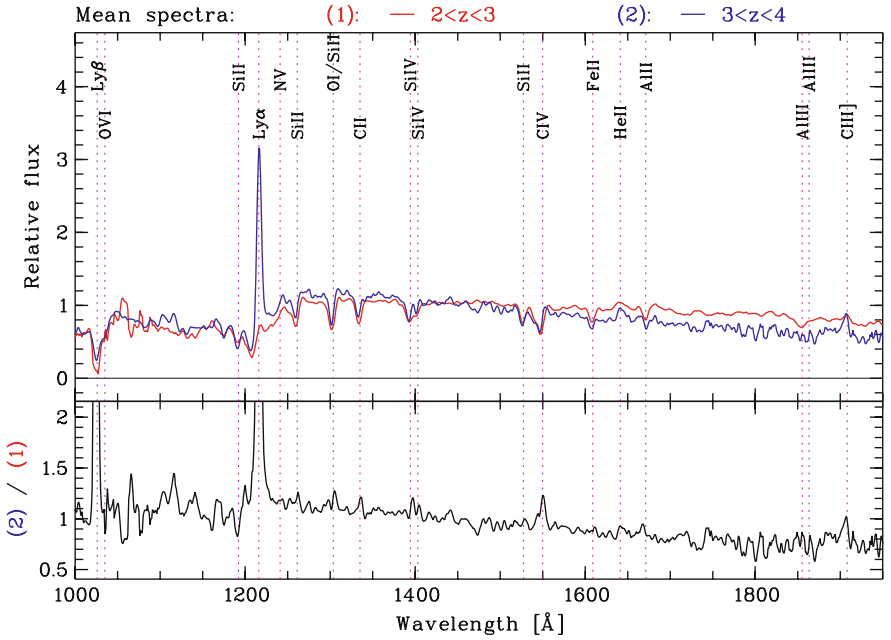

Fig. 7.19 Redshift evolution of the spectra of $2 < z < 4$ starburst galaxies. The *upper panel* shows composite spectra of galaxies with redshifts $2 < z < 3$ (spectrum (1), red line) and $3 < z < 4$ (spectrum (2), *blue line*). The lower panel shows the ratio of the two spectra, derived by dividing the higher-redshift spectrum by the lower-redshift one. Maxima in the ratio spectrum indicate stronger emission or weaker absorption at higher redshifts. (From Noll et al. [377])

over, the He II 1641 Å emission and the strength of the C III] intercombination line increase with increasing redshift. Since these two lines are very sensitive to the temperature of the line-emitting gas, the increased strength indicates a higher temperature of the interstellar gas and radiation field of the higher-redshift objects. Figure 7.19 also shows that the higher-redshift starburst galaxies have systematically steeper UV continua. Also evident in Fig. 7.19 is a larger continuum break at Lyα, which is consistent with the increase of the Lyman forest absorption with the redshift (see Sect. 6.9).

At least part of the observed redshift dependence of the observed spectral properties can be explained as an effect of the chemical evolution of the high-redshift starburst galaxies, which will be discussed in the next subsection. But other causes (such as a redshift dependent IMF) cannot be excluded. (For a more detailed discussion see, e.g., [377].)

7.1.2.9 Observed Chemical Abundances

As explained in Sect. 6.7, a reliable derivation of the chemical composition of the stars or of the interstellar matter of a galaxy requires high-quality spectroscopic observations. With the visual and infrared instrumentation available at present, spectra of sufficient resolution and adequate S/N can be obtained for starburst galaxies with redshifts up to about $z = 5$. However, with increasing redshift taking such spectra becomes more difficult and time consuming.

Among the most promising current projects aimed at deriving the chemical compositions at high redshifts is the international "Assessing the Mass-Abundance z Evolution" (AMAZE) program [330]. The objective of this project is to measure the chemical compositions of about 30 representative starburst galaxies in the redshift range $3 < z < 5$ using ground-based NIR spectroscopy. The abundances are determined by means of H II–region emission line ratios (Sect. 6.7). In detail, AMAZE uses the lines of [O II], [O III], [Ne III], and Hβ. At smaller redshifts Erb et al. [159] used the N II/Hα ratio (see Fig. 6.6) to derive the abundances of 87 star-forming galaxies at $z \approx 2.2$. Figure 7.20 shows first results from the AMAZE project, together with the results of Erb et al. and corresponding data for the lower-redshift universe.

The observations show that at least up to $z = 2.2$ the star-forming galaxies follow a well-defined "mass-metallicity" relation , as observed in the local universe (Sect. 2.2.3, Fig. 2.13). But Fig. 7.20 also shows that with increasing redshift the mass-metallicity relation is shifted toward lower abundances. Obviously, this means that the oxygen abundance of galaxies of a given stellar mass has been increasing steadily since the epoch corresponding to the redshift $z = 3.3$. The evolution of the chemical composition with cosmic time following from Fig. 7.20 is plotted in Fig. 7.21. According to this figure, in massive galaxies most of the chemical enrichment took place before the epoch corresponding to $z = 2$, and relatively little further enrichment occurred afterward. On the other hand, in the low-mass galaxies much of the chemical enrichment took place relatively recently at epochs corresponding to the redshift range $0.7 < z < 2.2$. This result is good agreement with the dependence

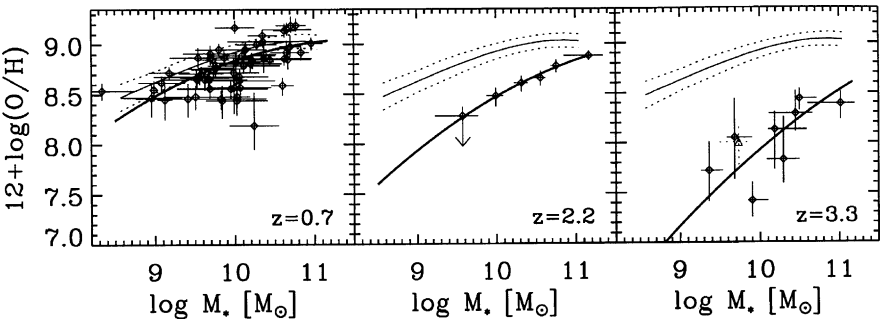

Fig. 7.20 Oxygen abundance as a function of the stellar mass of star-forming galaxies at different redshifts. The data for the redshifts 0.7, 2.2, and 3.3 are taken from, respectively, [450], [159], and [330]. The *thick solid lines* are fits to the data points. The *thin solid lines* and the *broken lines* represent the local ($z = 0$) mass-metallicity relation derived in [523] and reproduced in Fig. 2.13. (From Maiolino et al. [330])

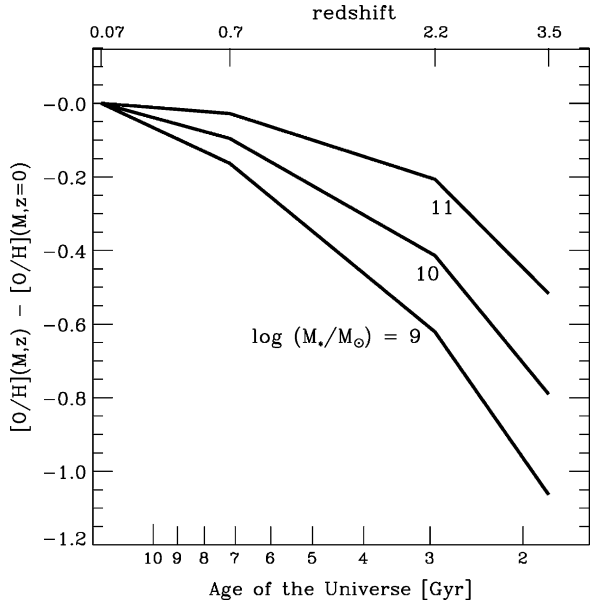

Fig. 7.21 Evolution of the oxygen abundance as a function of cosmic age for star-forming galaxies of different stellar masses, as inferred from the data plotted in Fig. 7.20. [O/H] stands for the logarithm of the ratio of the oxygen and hydrogen abundance, normalized to the solar value. (From Maiolino et al. [330])

on the mass and the cosmic time of the evolution of the cosmic star-formation rates described earlier in this section. Thus, the evolution of the chemical enrichment shows another aspect of the "downsizing" effect.

A similar redshift evolution of the chemical content of starburst galaxies is suggested independently by the observed absorption strength of the C IV resonance

doublet (e.g., [345]). In starburst galaxies this spectral feature normally is dominated by the stellar-wind absorption profiles of the very hot stars. Since the stellar winds of hot stars are driven by radiation pressure on ions such as C IV, these winds depend critically on the metallicity. As a result, the C IV 1550 Å absorption equivalent widths can be used as an indicator of the chemical abundances of starburst galaxies (see Sect. 6.7).

Figure 7.22 shows the mean C IV equivalent width of a sample of local starburst galaxies ($z = 0$) and of samples of high-redshift starburst galaxies as a function of redshift. As expected from the abundance evolution observed in the H II region emission lines, the C IV equivalent width decreases with redshift. A quantitative calibration of Fig. 7.22 in terms of abundance evolution is complicated by the fact that the high-redshift starburst galaxies of Fig. 7.22 are on average brighter and more massive than the galaxies of the local sample. If we compare in Fig. 7.22 only the equivalent widths of galaxies of similar brightness, and if we assume the empirical calibration of the C IV absorption equivalent width suggested by Mehlert et al. [345] (Eq. 6.11), we get from Fig. 7.22 a metallicity increase of about a factor 2.2 between $z \approx 2.3$ and $z = 0$, and an increase of a factor ≈ 2.9 between $z \approx 3.3$ and $z \approx 2.3$, which is in reasonable agreement with the abundance evolution of the massive starburst galaxies derived from interstellar emission lines and documented in Fig. 7.21.

The offset of the lines connecting the data points for the bright galaxies (triangles) and the fainter galaxies (squares) in Fig. 7.22 also seems to confirm qualitatively that the brighter, more massive galaxies experience their chemical enrichment at earlier epochs.

As discussed in Sect. 6.7 an independent estimate of the redshift evolution of the metallicity of high-redshift galaxies can be obtained from the photospheric lines and blends of the galaxy spectra (Fig. 6.7). This approach requires spectra

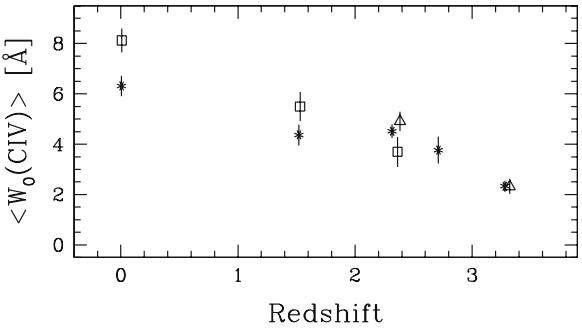

Fig. 7.22 Equivalent width of the C IV 1550 Å resonance doublet as a function of the redshift for a sample of high-redshift star-burst galaxies [345] and a sample of local ($z \approx 0$) starburst galaxies [225]. The *asterisks* represent the total samples. The *open triangles* correspond to galaxies brighter than $M_B = -22.3$. The *open squares* correspond to galaxies $-21.5 < M_B < 20.4$. The *error bars* indicate statistical mean errors resulting from the scatter within the corresponding samples. (From Mehlert et al. [345])

of medium or high spectral resolution and good S/N. Modern spectrographs at 8 to 10-m telescopes make it possible to obtain such spectra of high-redshift galaxies. However, these observations require long integration times and, consequently, substantial amounts of observing time at large instruments. Therefore, so far, medium and high-resolution spectroscopic studies of high-redshift galaxies have been limited to relatively few objects. Moreover, in many cases the available exposure times were too short to obtain high-quality results. Finally, as pointed out in Sect. 6.7, calibrations with different synthetic spectra tend to result in different absolute abundances. Nevertheless, the limited results on the metallicity sensitive photospheric features appear (within the error limits) consistent with the data derived from interstellar emission lines and the C IV stellar-wind features. This is demonstrated by Fig. 7.23. In this figure the equivalent widths of two abundance-sensitive absorption features in high-redshift starburst galaxies are plotted as a function of the redshift. Although the scatter of the individual measurements is large, the data support a significant increase of the metallicity of the galaxies of the observed sample between $z \approx 3.4$ and $z \approx 2.5$. With about a factor 4 the increase indicated in Fig. 7.23 is somewhat larger, but within the error limits still in agreement with the values derived from the H II region emission lines for massive galaxies (Fig. 7.21). On the other hand, a more recent study of a larger sample of photospheric spectra of star-forming galaxies resulted in somewhat lower metalicities at $z \approx 2$ [217]. A possible cause of these differences could be the systematic errors in the calibration with synthetic spectra (e.g., [347]).

In principle, chemical abundances can also be derived by analyzing the interstellar absorption lines of the high-redshift galaxies. However, since these lines normally are optically thick, their strength depends not only on the abundance, but

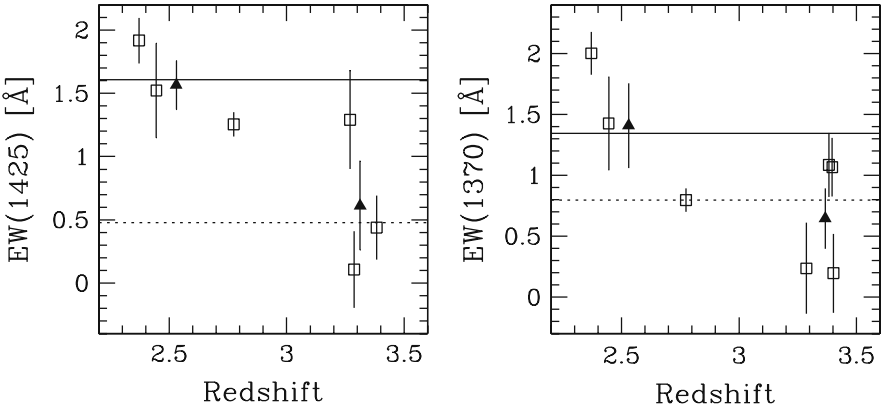

Fig. 7.23 Equivalent widths of the chemical-abundance sensitive 1425 Å and 1370 Å blends (Sect. 6.7) in the spectra of starburst galaxies as a function of redshift. The *open squares* represent individual measurements. The *filled triangles* represent mean values. The solid horizontal lines indicate the equivalents widths measured in synthetic spectra assuming a solar chemical composition ($Z = Z_{\odot}$), the broken lines indicate the corresponding values for $0.25Z_{\odot}$. (From Mehlert et al. [347])

also on the velocity fields of the gas, and on the covering factor of the absorbing clouds. In order to disentangle these effects, spectra of higher resolution and better S/N are needed than for the analysis of abundance sensitive photospheric features. Most high-redshift galaxies are too faint to obtain such spectra. An exception are gravitationally lensed galaxies with high magnification factors. An example is the $z = 2.728$ galaxy MS 1512-cB58, which has been investigated in detail by Pettini et al. [415, 416]. For this galaxy it was possible to derive in addition to the overall chemical abundance ($Z \approx 0.4 Z_{\odot}$) also abundances for individual chemical elements. Results of this study are reproduced in Fig. 7.24.

When interpreting measurements of the chemical composition of the interstellar gas, the fraction of the matter condensed into dust grains has to be taken into account (see Sect. 2.2.3). Therefore, instead of comparing the absolute abundances derived from interstellar absorption lines with the solar abundances, it is more meaningful to make a comparison with the abundances found in the diffuse interstellar medium of the Milky Way galaxy, given in Fig. 7.25. As shown by comparing Figs. 7.24 and 7.25, in the ISM of the high-redshift starburst galaxy MS 1512-cB58 at least nitrogen and magnesium do not follow the relative abundance pattern known from our own galaxy. As discussed in detail in [415], these differences (in particular the

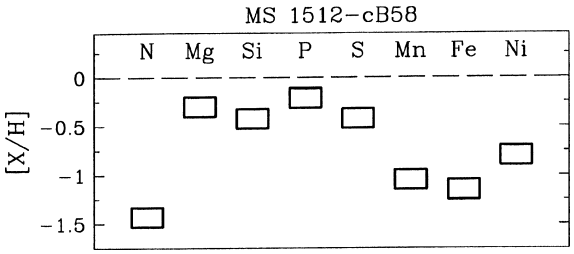

Fig. 7.24 Chemical abundances relative to the solar values of individual elements derived from interstellar absorption lines in the spectrum of the $z = 2.728$ starburst galaxy MS 1512-cB58. The ordinate gives the logarithm of the ratio of the abundance of the element in question and the abundance of hydrogen, normalized to the solar value. The height of the boxes indicate the statistical uncertainties of the measurements. (From Pettini et al. [415])

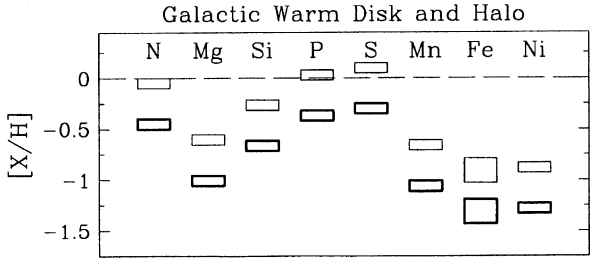

Fig. 7.25 Chemical abundances relative to the solar values of the elements of Fig. 7.24 in the diffuse interstellar clouds of the Milky Way galaxy (*thin-line boxes*). To facilitate a comparison with Fig. 7.24, the *thick-lined boxes* show the same abundance pattern shifted down by 0.4 dex. (From Pettini et al. [415])

small nitrogen abundance) seems to indicate that the chemical enrichment of MS 1512-cB58 took place by means of supernovae of type II within the last about 300 Myr before the observed light was emitted.

7.1.2.10 Components and Kinematics of the Interstellar Gas

Like the local galaxies, the high-redshift star-forming galaxies contain an interstellar medium which consist of several distinct components. Some of these components can be directly observed. As noted already, the gas which is ionized by the UV radiation of the hot young stars can be studied by analyzing the emission line spectra of these galaxies. Information on the cool, neutral gas can be obtained by observing the absorption lines of atoms and ions which are located in front of the observed stars. Finally, the cold, molecular gas can be studied by means of the low-excitation molecular transitions observable at submm and radio wavelengths (e.g., [28] and [112]).

The interstellar emission line spectra of the high-redshift starburst galaxies have already been discussed in the context of the chemical abundance derivation. The relative line strengths in these spectra usually are similar to those observed in local starburst galaxies. They normally can be reproduced well by photoionization models assuming that the ionization is due to the UV radiation of the hot stars which produce the observed UV continua. Deviations observed in individual spectra can be explained by AGN contributions to the radiation field or by (compared to typical local galaxies) somewhat higher electron densities and temperatures (e.g., [318]).

The spectral lines emitted by the low-density interstellar gas are intrinsically very narrow. But the profiles which are observed in high-redshift starburst galaxies have Doppler widths which typically are larger than the velocity of sound of the interstellar plasma. These large line widths result from the large-scale motions of the gas. Hence the line profiles can been used to derive (or at least to constrain) the velocity fields of the gas.

Hα and [O III] line profiles of more than 100 distant star-forming galaxies have been studied by Erb and colleagues [160], who found for two sub-samples with mean redshifts of $z = 2.2$ and $z = 3.1$ mean velocity dispersions of, respectively, (108 ± 41) km s^{-1} and (82 ± 28) km s^{-1}. Of the observed individual velocity dispersions of the line emitting gas 98% were found to be between 40 km s^{-1} and 200 km s^{-1}. From their data Erb et al. estimated for the $z \approx 2$ galaxies in their sample a mean dynamical mass of $7 \times 10^{10}\ M_{\odot}$ inside a radius of about 6 kpc.

Another important source of information on the structure and the kinematics of the interstellar medium in high-redshift starburst galaxies is the analysis of the profiles of the interstellar absorption lines. Particularly valuable have been observations of the absorption lines of atoms and ions with low ionization potentials (referred to as "low-ionization interstellar lines" (LIS) lines). As shown in Fig. 7.13 low-ionization absorption lines, such as O I 1302.2 Å, C II 1334.5 Å, N I 1199.9 Å, N II 1741.5 Å, Si II 1526.7 Å, Al II 1670.8 Å, Al III 1854.7 Å, 1862.8 Å, and Fe II 1608.5 Å are among the most conspicuous features of the UV spectra of high-redshift starburst galaxies. Since the corresponding low-ionization stages do

not survive in the atmospheres of the hot stars which produce the UV continuum, these absorption lines must originate in cool, essentially neutral, volumes of the interstellar medium in front of the continuum sources. Thus, these lines provide selective information on the cool and neutral components of the interstellar gas.

While low-ionization interstellar absorption lines are observed in all spectra of high-redshift starburst galaxies, their strength varies greatly between different objects of the same redshift. This is illustrated by Figs. 7.26 and 7.27. Particularly large equivalent widths of the low-ionization absorption lines are measured in the gravitationally lensed $z = 2.272$ starburst galaxy MS 1512-cB58 (Fig. 7.26). Since the flux from this galaxy is amplified by a factor of ≈ 30, it is the most thoroughly investigated galaxy in this redshift range. A listing of the extensive literature on this important object can be found in [453].

While the photospheric line spectrum of MS 1512-cB58 is rather typical, for a starburst galaxy at this redshift, its interstellar absorption lines are stronger than average. Compared to the composite spectrum of $z \approx 3$ starburst galaxies reproduced in Fig. 7.13, the spectrum of MS 1512-cB58 shows equivalent widths of the low-ionization interstellar lines which are larger by a factor of about 1.7. In the spectrum of FDF-5903, reproduced in the upper panel of Fig. 7.27, the interstellar lines are almost as strong (≈90%) as those of MS 1512-cB58. On the other hand, the interstellar absorption lines of FDF-6024 (lower panel of Fig. 7.27) are much

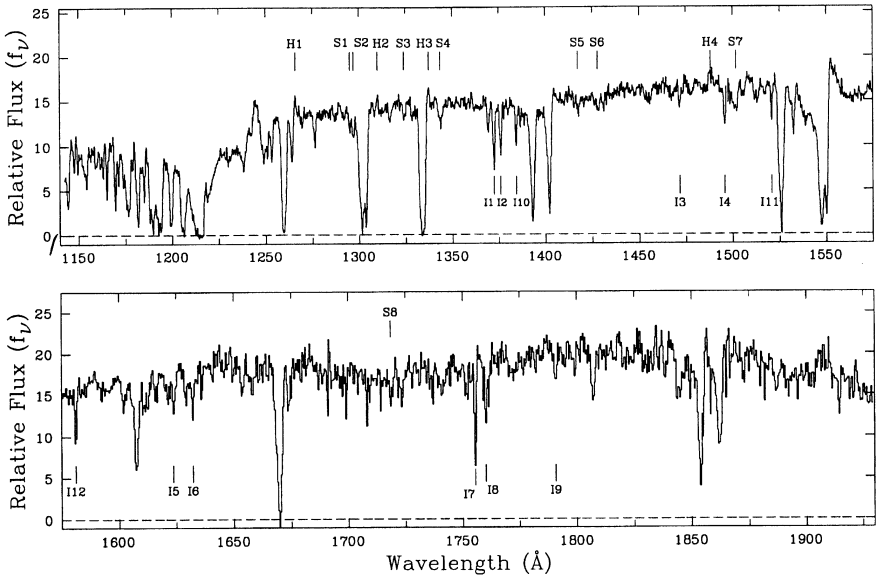

Fig. 7.26 Medium-resolution ($1700 < R < 3100$) spectrum of the $z = 2.727$ starburst galaxy MS 1512-cB58. The (weak) emission lines indicated and labeled "H" are ascribed to H II region emission. Lines originating in the stellar photospheres are labeled "S." The *lines* labeled "I" are of interstellar origin. The *strong lines* not marked in the figure can be identified by comparison with Fig. 7.27. (From Pettini et al. [416])

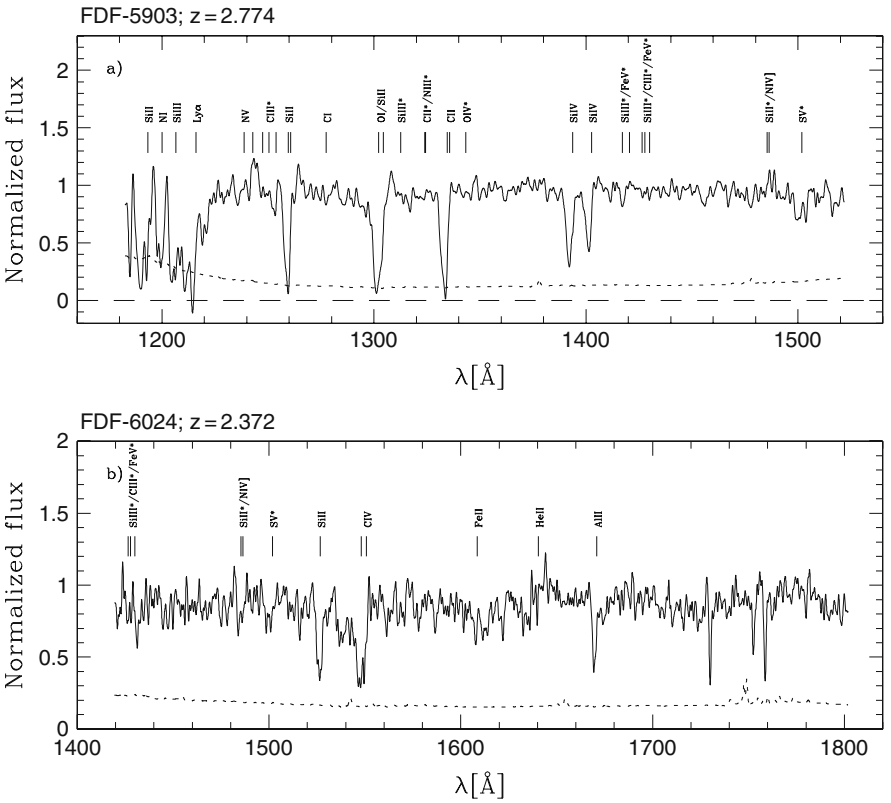

Fig. 7.27 Medium-resolution ($R = 2100$) spectra of the two star-burst galaxies FDF-5903 ($z = 2.774$) and FDF-6024 ($z = 2.372$). The *dotted lines* indicate the noise level. While the profiles of the strongest interstellar absorption lines in the spectrum of FDF-5903 show zero rest intensity in the line centers, the interstellar lines in the spectrum of FDF-6024 have significant rest intensities. The three strong absorption features at 1720-1760 Å in the spectrum of FDF-6024 are Fe II absorption lines of a foreground IGM absorption system at $z = 1.4885$. (From [347])

weaker than in MS 1512-cB58. Lines, such as Si II 1526.7 Å and Al II 1670.8 Å, which have saturated ("black") cores in MS 1512-cB58, show significant rest intensities in FDF-6024. In all the spectra which are compared above the profiles are sufficiently well resolved, to exclude the possibility that the differences are due to resolution effects.

A clue to the origin of the strength differences of the LIS can be derived from a comparison of lines with different oscillator strengths. Under optically thin conditions the equivalent widths of the interstellar resonance lines are expected to be proportional to the oscillator strengths. If the lines are fully saturated, the absorption becomes independent of the line parameters. The analysis of the galaxy spectra shows that the cores of the strong LIS lines in distant starburst galaxies are saturated (e.g., [347]). Thus, the observed non-zero rest intensities in part of the high-redshift starburst galaxies can be explained only by assuming that part of the

continuum radiation reaches us directly with no (or only little) line absorption along the line-of-sight, while another part suffers complete absorption of the center of the strong LIS lines. On first approximation the ratio between the absorbed and the unaffected (or little affected) flux is proportional to the surface fraction of the continuum source which is covered by cool, line absorbing clouds. Therefore, this ratio is called the "covering factor." For MS 1512-cB58 the covering factor has been estimated to be about 98% [226]. Typical covering factors for UV-bright high-redshift star burst galaxies are of the order 50–70%.

Different lines in the spectra of high-redshift galaxies often show different radial velocities. Identical redshifts are normally observed for the stellar (photospheric) absorption lines and for most of the interstellar emission lines. Therefore, these lines are assumed to represent the systemic velocity of the galaxies. In most high-redshift starburst galaxies the low-ionization interstellar absorption lines are blueshifted (by up to several 10^2 km s^{-1}) relative to the stellar photospheric spectra, while the Lyα emission lines often appear redshifted relative to the systemic velocity (e.g., [4, 347, 415, 416, 471]). The blueshift of the interstellar lines suggests an outward flow of (at least part of) the observed cool gas. The redshift of the Lyα lines appears to be due to interstellar and intergalactic absorption in the blue wing and a higher escape probability of Lyα photons in the red wing of the line. Because of the different sign of the line shifts of Lyα and the LIS lines, particularly large relative shifts are observed between the low-ionization absorption lines and the Lyα emission. The distribution of the corresponding velocity difference in the spectra of $z \approx 3$ starburst galaxies is given in Fig. 7.28. As shown by Shapley et al. [471] this

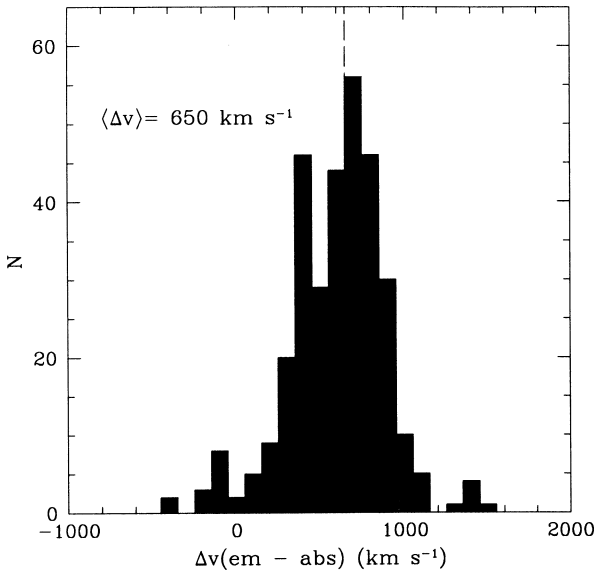

Fig. 7.28 Distribution of the radial velocity differences of the Lyα emission peaks and the low-ionization interstellar absorption lines. The *broken line* indicates the median of the distribution. (From Shapley et al. [471])

velocity difference correlates with the Lyα equivalent widths. The largest velocity differences (about 800 km s^{-1}) are observed for galaxies with weak Lyα emission, while at large Lyα emission equivalent widths the velocity difference approaches a lower limit of 400 km s^{-1}.

The profiles of the low-ionization interstellar absorption lines have been found to be on average broader than the emission line profiles. Independent spectral samples give similar mean FWHM values between 500 km s^{-1} and 600 km s^{-1} (e.g., [347, 471]). Part of the line width difference between interstellar emission and absorption lines may be due to different properties of the samples in which the lines were measured. A physical explanation for the difference could be that the emission line widths reflect mainly the orbital motion of the interstellar matter (i.e., mainly the galaxies' rotation) while the absorption line widths are due to a combination of the orbital motions and outflows of the cool gas in the outer regions of the galaxies.

Examples of the observed line profiles of LIS lines are presented in Figs. 7.29 and 7.30. As shown by these figures, some galaxies show essentially symmetric profiles

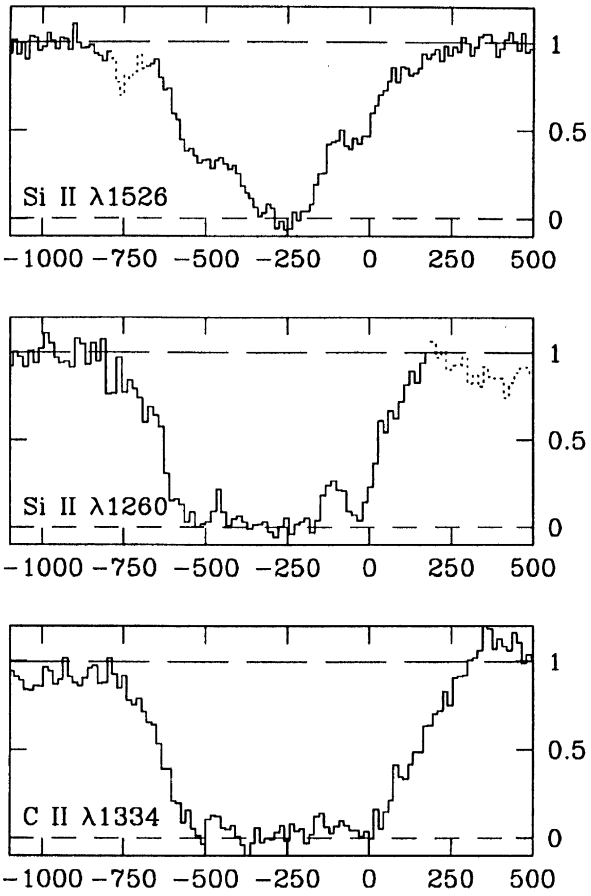

Fig. 7.29 Profiles of interstellar absorption lines of Si II and C II in the spectrum of the $z = 2.728$ starburst galaxy MS 1512-cB58. The individual plots present the normalized spectral flux F_λ as a function of the Doppler velocity relative to the galaxy's systemic velocity in km s^{-1}. (From Pettini et al. [415])

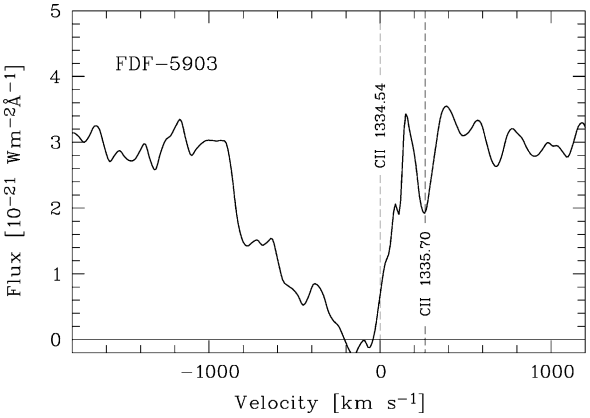

Fig. 7.30 Profile of the interstellar C II 1334 Å absorption line in the spectrum of the $z = 2.774$ starburst galaxy FDF-5903. (From Mehlert et al. [347])

with a blueshifted line center (Fig. 7.29), while other galaxies exhibit complex line profiles with extended blue wings (Fig. 7.30).

As noted already, the blueshifts of the interstellar absorption lines are normally ascribed to galactic winds. Spherical outflow geometries (with expanding neutral gas shells) and conical, bipolar outflows (as observed in some local starburst galaxies) have been suggested for these winds. In such models the differences of the observed line profiles can be explained as a result of different outflow geometries and different orientation of the flow directions relative to the line-of-sight. As in local starburst galaxies, the galactic winds are expected to be driven by the momenta of stellar winds, stellar radiation pressure effects, and of expanding supernova shells (e.g., [123, 339, 515, 516]).

While some conclusions on the internal kinematics and dynamics of distant starburst galaxies can be obtained from the emission and absorption line profiles in the integrated galaxy spectra, detailed information on the velocity fields of the interstellar gas can be derived only from spatially resolved spectroscopic observations. Because of the small angular size of most high-redshift galaxies, the most efficient technique for obtaining such data is a combination of integral field spectroscopy and adaptive optics (e.g., [61, 176, 201, 302]). Thus far, this technique has been applied to starburst galaxies in the redshift range $1.4 < z < 3.4$, and spatial resolutions between 1 kpc and 5 kpc have been achieved. An example is given in Fig. 7.31. Some of the observed starburst galaxies show only random velocity fields (e.g., [302]), while others show kinematic evidence for merger processes [371]. But part of the observed high-redshift starburst galaxies were found to have systematic velocity gradients which appear similar to those observed in local disk galaxies (e.g., [61, 201]). For those high-redshift starburst galaxies which were selected on the basis of (rest-frame) UV and visual colors, the inferred circular velocities and disk radii were found to be surprisingly similar to those observed in local Sb/c galaxies. However, according to Bouché and colleagues [61], the expected redshift

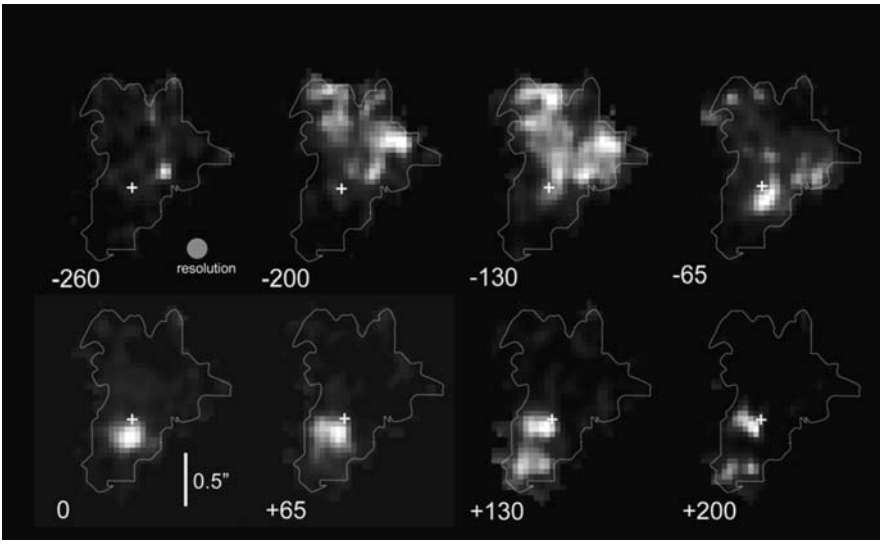

Fig. 7.31 Hα images of the $z = 2.38$ starburst galaxy BzK-15504 obtained with the SINFONI integral field spectrograph at the ESO VLT in adaptive-optics mode. The image labeled "0" shows the Hα flux emitted at the systemic velocity of the galaxy. The images labeled -260, -200, ... ,+ 200 show the Hα emission at the corresponding velocity shift (in km s^{-1}) relative to the systemic velocity. The Hα emitting gas in the upper-right part of the galaxy is approaching us, while the lower-left part is receding. This is explained by a rotation of the galaxy. The patchy structure of the emission appears to be due to a non-uniform distribution of the star-formation activity. Such patterns are rather typical for starburst galaxies. The spatial resolution of the images is about 0.15" or 1.2 kpc. For details of the observations and their interpretation see [201]. (Courtesy European Southern Observatory. ESO Press Photo 31b/06)

dependence of the relation between disk radii and the circular velocity (Eq. 3.45) could not be detected in the observed samples. Only observations of larger, more representative samples can clarify whether this unexpected result is due to selection effects, or due to incorrect model assumptions. The dynamical masses derived from the high-resolution IFU studies are in the range $10^9 < M/M_\odot < 10^{12}$.

The spatially resolved observations of the distant star burst galaxies have also been used to estimate gas surface densities of the observed apparent disks [61]. For the same galaxies SFRs could be derived from the UV flux, emission lines or from the FIR flux. Thus, the observed surface densities could be compared with the empirical Schmidt and Kennicutt SFR laws (Eqs. 2.10 and 2.11). Interestingly, the high-redshift starburst galaxies (including submm-selected objects) and the local star-forming disks seem to follow the same empirical relations between the SFR and the gas surface density. Moreover, the observed relation seems to remain unchanged over many decades of the SFR (e.g., Fig. 3 of [61]).

Apart from using high-angular resolution integral-field spectroscopy, spatially resolved information on the dynamics of distant galaxies can be derived by reconstructing the intrinsic light distribution of lensed galaxy images. Since the main

effect of strong gravitational lensing is a redistribution of the light from a distant galaxy over a larger angular area (without reducing the surface brightness), gravitational lensing improves the angular resolution. On the other hand, because of the complex optical aberrations of gravitational lenses (see [458]), a reliable knowledge of the lens properties is required for a reconstruction of the original galaxy image. A reliable reconstruction is possible, if a sufficient number of lensed objects can be evaluated.

Among the successful studies of the internal kinematics of a lensed high-redshift starburst galaxy is an investigation of the prominent arc image visible in Fig. 5.9 (Sect.5.2.4). Nesvadba and colleagues [372] succeeded in deconvolving the lensed image into an inner core and the outer region of the lensed galaxy, reaching a spatial resolution of about 200 pc. Using this data, Nesvadba et al. were able to derive the rotation curve and a dynamical mass of the galaxy and of its dark-matter halo ($\approx 5 \times 10^{11}$ $M_{\odot}$).

7.1.2.11 Interstellar Dust in High-Redshift Starburst Galaxies

Evidence for the presence of interstellar dust in high-redshift galaxies was found when the observed UV continua were compared to synthetic spectra showing the same stellar absorption lines. Due to the presence of dust, the UV and optical continua are attenuated and reddened. As a result of the absorption of UV photons, the dust grains are heated to a temperature where the thermal emission just balances the absorption of UV radiation. As explained in Sect. 2.2.5, the equilibrium temperatures are of the order 10 K–10^2 K. Thus, the dust re-emits the absorbed UV radiation at FIR wavelengths.

In Sects. 6.3 and 6.8 the UV continuum slope parameter β (Eq. 6.4) has been introduced. As pointed out in these sections, β is a convenient measure of the dust absorption in high-redshift starburst galaxies. If no dust absorption is present, young high-redshift starburst galaxies have $\beta \approx -2.5$. With increasing dust absorption β becomes less negative.

If the dust distribution in the galaxies is spherically symmetric and if the FIR radiation of the dust is emitted isotropically, the amount of light absorption (inferred from β) and the FIR luminosity of the dust should be correlated. In fact, for local starburst galaxies an empirical relation between β and the FIR luminosity has been well established [351]. Since the dust emission of most UV-bright high-redshift starburst galaxies is below or close to the detection limit of present mm-wave telescopes (e.g., [119]), our knowledge on the the thermal dust emission of $z > 2$ galaxies is limited. Good data exist for the lensed UV-bright starburst galaxy, MS 1512-cB58 ($z = 2.727$). Unexpectedly this galaxy was found to emit only about 25% of the (redshifted) FIR luminosity predicted from the observed β and the $\beta - F_{FIR}$ relation observed for local starburst galaxies [27]. This seems to indicate that the assumptions listed above are not correct, or that the dust properties are different at this redshift.

A conspicuous property of the dust absorption, as characterized by the exponent β, is a tight correlation between β and the strength of the low-ionization interstellar

absorption lines in starburst galaxies (e.g., [225, 377, 471]). As shown by Fig. 7.32, this correlation shows no detectable redshift dependence. Heckman et al. [225] suggested that the $W_{LIS} - \beta$ relation is caused by an increase of the interstellar turbulent velocity field with increasing dust content. The stronger interstellar turbulence could be related to a higher star-formation rate in the more dusty galaxies. An alternative (and simpler) interpretation was suggested by Shapley et al., who proposed that the variations in the dust absorption and the LIS line absorption are due to varying covering factors. As pointed out in the previous section, the covering factor measures the fraction of the surface of a galaxy which is seen through cool interstellar clouds. Since the dust is assumed to be concentrated in these cool clouds, a correlation of the type observed is to be expected.

As shown by Fig. 7.32, β is also correlated with the strength of the C IV absorption. In fact, the dependence is even stronger than in the case of the LIS lines. While W_{LIS} varies by a factor ≈ 2, $W_{\rm CIV}$ varies by a factor 4 in the same β range. This seems to rule out that the variation of the LIS lines *and* that of the C IV line can be caused by the covering factor alone. Moreover, while interstellar components contribute to the C IV line, in most starburst galaxies the C IV absorption is dominated by stellar-wind features. Finally, the C IV equivalent width is known to depend sensitively on the chemical composition (see Sect. 6.7). Since the formation of dust requires heavy elements, β is expected to depend on the chemical composition as well. Both the chemical composition and β are known to change with the cosmic age and redshift (see Fig. 7.21 and [377]). Therefore, it seems likely that at least part of the $W_{\rm CIV} - \beta$ dependence (and possibly part of the $W_{LIS} - \beta$ dependence) is due to the effects of the metallicity on these parameters.

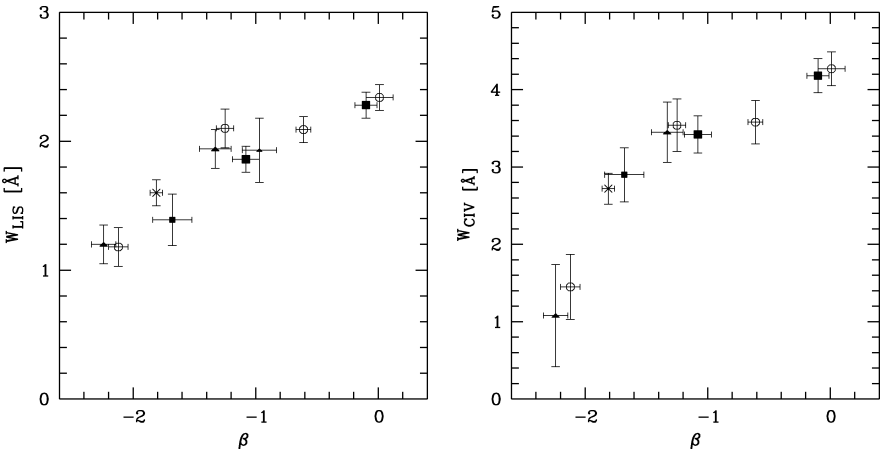

Fig. 7.32 The average equivalent widths of the low-ionization interstellar lines (as defined in the caption of Fig. 7.18) and of the C IV 1550 Å blend in the spectra of $z \approx 3$ starburst galaxies as a function of the reddening-sensitive UV continuum slope parameter β. The *open circles* represent the total sample of high-redshift galaxies used for this study. The *filled squares* and *filled triangles* refer to the galaxies in the redshift ranges $2 < z < 3$ and $3 < z < 4$, respectively. (From Noll et al. [377])

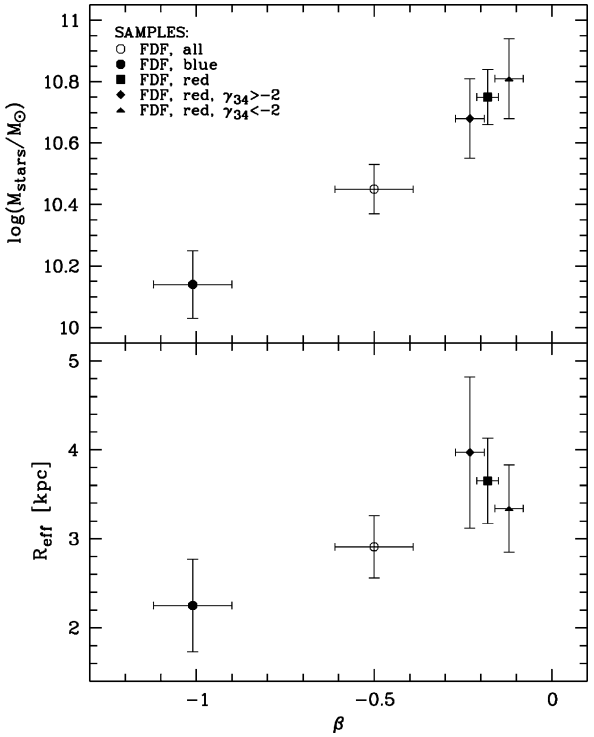

Fig. 7.33 Empirical relations between the dust reddening parameter β and the stellar mass and effective radius of UV-bright starburst galaxies in the redshift range $2 < z < 2.5$. (From Noll and Pierini [378], where details are described)

Figure 7.33 indicates that the dust parameter β also correlates with the stellar mass and with the effective radii of the high-redshift starburst galaxies [378]. The more massive and the larger starburst galaxies tend to show more dust absorption than the smaller, low-mass objects. The lower internal dust absorption in small and low-mass galaxies may have a simple geometric origin. In small galaxies the geometric light paths are shorter and the chances of encountering a dust cloud is lower. On the other hand, the smaller reddening could also be the result of a lower dust content resulting from the (on average) smaller chemical enrichment of the low-mass objects.

A complication for the use of the dust-reddening parameter β for the derivation of the light attenuation is the fact that the UV reddening law is not completely uniform for high-redshift starburst galaxies. As a result, there exists no reliable relation between β and the amount of UV absorption (e.g., [378]). Variations of the dust reddening law are well known from observations in our Milky Way galaxy and from nearby galaxies. Among the evidence for variations of the dust reddening law is the behavior of a broad maximum of the interstellar extinction function at 218 nm. In the literature this local maximum in the dust extinction law is referred to

as the "UV bump." This feature is usually ascribed to graphitic carbon absorption of dust grains at these wavelength. The UV bump is a prominent feature of the dust reddening law observed in the Milky way galaxy. But the reddening law derived for local ($z \approx 0$) starburst galaxies does not show a UV bump.[1] A careful evaluation of the SEDs of high-redshift starburst galaxies shows that UV-bump absorption is present in part of the high-redshift objects.

S. Noll and colleagues [378, 379] suggested UV continuum slope parameters which are sensitive to different reddening laws. Among these slope parameters is the quantity γ_{34}, which is the difference of two continuum slope exponents measured at wavelengths blueward and redward of the UV bump. Thus, γ_{34} is a sensitive measure of possible UV-bump absorption. An analysis of the SEDs of various different samples shows that a significant fraction of the $1 < z < 2.5$ starburst galaxies is affected by UV-bump absorption. UV-bump absorption is observed most frequently in galaxies with a high amount of reddening. These galaxies seem to contain a mixture of dust grains with different physical properties. For such galaxies estimates of the dust extinction based on continuum slope parameter β and a normal Calzetti law may not be correct. Since there are also other indications that high-redshift galaxies can have different dust properties, corrections for dust absorptions have to be taken with caution. Fortunately, this problem is less critical for very high redshifts, since (probably as a result of the decreasing heavy element abundance) the dust reddening is known to decrease rapidly with redshift (see Sect. 7.1.4).

The UV-bump absorption also correlates with various other spectral properties. This includes an apparent correlation with the metallicity dependent C IV absorption strength. However, as shown in Fig. 7.34, the high-redshift star-burst galaxies with blue colors show (like the local starburst galaxies) positive values of γ_{34}, which are consistent with the absence of a UV bump in the extinction law. Highly negative values of γ_{34} and large values of W_{CIV} are observed for starburst galaxies with red colors only. This seems to indicate that the presence of a blue bump is related to the galaxy properties, but not directly dependent on the metallicity.

Since the total dust absorption of high-redshift galaxies depends on the galaxy mass (Fig. 7.33), γ_{34} has also been checked for a possible dependence on this parameter. No correlation between the mass and the UV-bump strength could be found [379].

7.1.3 Lyα Emission Galaxies (LAEs)

Although the high-redshift Lyα emitting galaxies are, strictly speaking, a subclass of the UV-bright starburst galaxies, they are discussed separately, since they show qualitatively distinct properties and physical parameters which differ from those of the majority of the UV-bright high-redshift objects.

---

[1] A reason for the absence of UV-bump absorption in the known samples of local starburst galaxies could be a selection effect. These local samples contain mainly galaxies with relatively low dust absorption, while the UV bump is known to occur preferentially in significantly absorbed objects.

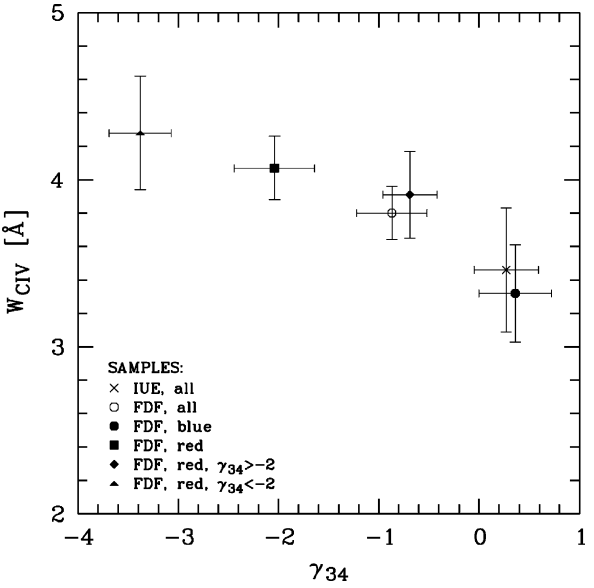

Fig. 7.34 Empirical relation between the C IV absorption strength and the UV-bump indicator γ_{34}. With increasing UV-bump absorption γ_{34} becomes more negative. The *asterisk* represents the sample of local starburst galaxies observed with the International Ultraviolet Explorer (IUE) satellite. The other symbols represent starburst galaxies with redshifts $2 < z < 2.5$ from the FDF survey. (From Noll and Pierini [378])

7.1.3.1 Discovery and Definition

As pointed out at the beginning of this chapter, in a pioneering paper published in 1967, R. B. Partridge and P. J. E. Peebles predicted the existence of high-redshift starburst galaxies with strong Lyα emission lines. But, as described in Sect. 7.1.2 most spectra of the first significant samples of high-redshift starburst galaxies showed the Lyα line in absorption or only weakly in emission. In view of the quenching of Lyα resulting from the combination of resonance scattering and dust absorption (explained in Sect. 7.1.2), the weakness of the Lyα emission in high-redshift starburst galaxies is not surprising. However, in 1996 Esther Hu and Lennox Cowie [236] actually found several Lyα emitting galaxies at $z = 4.55$. In a subsequent paper Hu et al. [237] reported the discovery of a "substantial population" of galaxies with conspicuous Lyα emission with redshifts of $z \approx 3.4$ and $z \approx 4.5$. These galaxies had Lyα equivalent widths as high as predicted in the 1967 paper of Partridge and Peebles ([401]). But their UV continuum flux was only about 1% of that predicted in 1967.

Today the objects discovered by Hu et al. are referred to as "Lyα emission galaxies," "Lyα galaxies," or "Lyman-alpha emitters" (LAEs). To distinguish them from normal UV-bright starburst galaxies, LAEs are usually defined as distant starburst

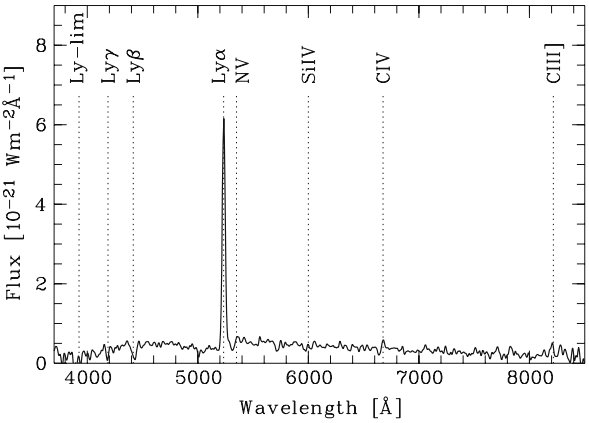

Fig. 7.35 Example of a spectrum of a Lyα emission galaxy. This $z = 3.304$ galaxy has a Lyα rest-frame emission equivalent width of 79 Å. (From Tapken et al. [513])

galaxies having rest-frame Lyα emission equivalent widths > 20 Å[2]. Some LAEs show equivalent widths > 250 Å. Their UV continuum luminosities typically are relatively low. Among continuum-flux limited samples of high-redshift starburst galaxies the LAEs are on average fainter than galaxies without strong Lyα emission (e.g., [196]). An example of an LAE spectrum is shown in Fig. 7.35. From the observed large Lyα equivalent widths, it is clear that the observed low-continuum luminosities cannot be caused by dust absorption. Obviously, these galaxies are intrinsically faint. But the relatively high space density of the distant Lyα galaxies derived by Hu et al. [237] indicated that the LAEs provide major contributions to the star formation and to the cosmic UV radiation field.

In the decade following the discovery of Hu et al., many additional LAEs with redshifts $3.4 < z < 6.5$ have been found and studied (see Sect. 7.1.1 and the references listed in the caption of Fig. 7.36). Most of these galaxies were detected during line-emission searches of the kind described in Sect. 5.1.3. But LAEs were also discovered during spectroscopic follow-up observations of UV-continuum selected high-redshift galaxies, and some were detected independently by both methods.

7.1.3.2 Frequency and Lyα Luminosity Function of the LAEs

In view of the contribution of LAEs to the UV radiation field, many papers have been devoted to the space density and the luminosity function of these objects. The different derivations of the LAE luminosity function have been discussed by Ouchi

[2] Note that different authors use somewhat different definitions. While there is agreement that LAEs are objects with Lyα emission equivalent widths > 20 Å, some authors regard only the galaxies detected in narrow-band searches (but not emission line galaxies detected with other techniques) as LAEs. Other authors include AGN detected in photometric searches in the LAE definition.

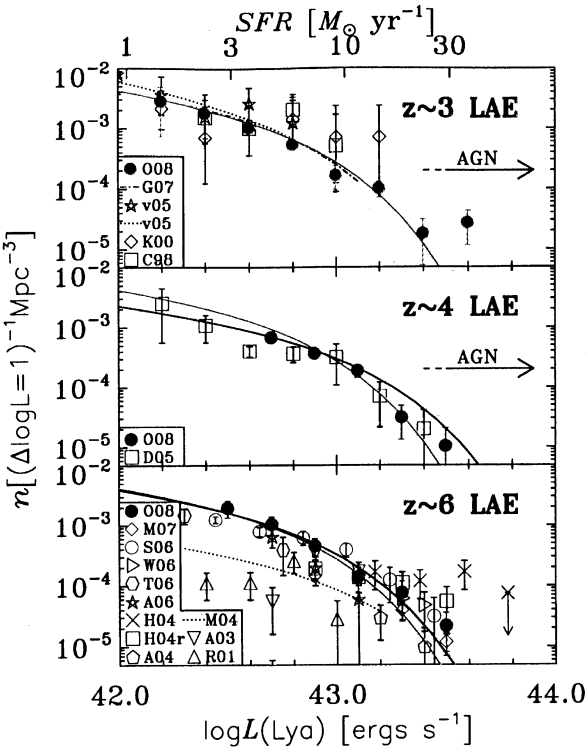

Fig. 7.36 Luminosity functions of the Lyα emitters at different redshifts. The sources of the individual data points are given by the reference codes on the left margin of the figures. Each code consists of a letter (indicating the first author) and a two-digit number (indicating the year of publication). The corresponding numbers in the list of references at the end of this book are A06: [9], A03: [10], A4: [11], C98: [114], D05: [124], G07: [212], H04: [238], H04r: [238], recalculated in [394], K00: [288], M04: [333], M07: [366], O08: [394], R01: [432], S06: [475], T06: [511], v05: [535], W06: [554]. (From Ouchi et al. [394])

and colleagues [394] and are summarized in Fig. 7.36. As shown by this figure the data published after about 2004 are fairly consistent and constrain the LF of the LAEs rather well. Like other LFs, the Lyα luminosity function of the LAEs can be well approximated by a Schechter function. Figure 7.36 also shows that the Lyα luminosity function does not evolve significantly with the redshift. At least in the redshift range $3 < z < 6$ the intrinsic LF appears to stay essentially constant, and no statistically significant variations of the Schechter parameters with the redshift could be detected. According to the analysis of Ouchi et al., the mean characteristic Lyα luminosity is of the order $L^*_{Ly\alpha} = 8 \times 10^{42}$ erg s^{-1} ($= 8 \times 10^{35}$ W). The normalization was found to be about $\Phi^* \approx 5 \times 10^{-4}$ Mpc^{-3}. (The Schechter parameter α is not well constrained).

Less well known than the LF is the fraction of the LAEs among the UV-bright starburst galaxies. For several reasons this fraction is expected to vary with the

redshift. Since galaxy formation started from matter containing essentially only hydrogen and helium, which cannot form solids under astrophysical conditions, Lyα quenching by dust absorption should be less efficient at the highest redshifts. This could result in an increase of the LAE fraction with redshift. On the other hand, since the number of the "Lyman-forest" absorbers increases with z (see Sect. 6.9), the chances of IGM absorption in the blue wing (and possibly even in the red wing) of the LAE emission lines become larger with increasing redshift. This effect is expected to cause a decrease of the apparent LAE fraction with increasing redshift.

At very high redshifts the Lyman-forest absorption unavoidably also results in a reduction of the observed mean Lyα equivalent width relative to the intrinsic values. On the other hand, the observed maximum values are probably not much affected since, statistically, the LAEs with the largest apparent emission strengths are expected to be found among the objects with low IGM absorption.

For local starburst galaxies the fraction of LAEs is not well known, since the Lyα emission of the extended Earth atmosphere (the "geocorona") hinders Lyα observations of $z \approx 0$ galaxies. For $z \approx 2$ Reddy and colleagues [429] find a LAE fraction of 8%. Rather well established is the LAE fraction at $z \approx 3$. For this redshift there exist large and well-defined samples of high-quality galaxy spectra. According to these samples about 25% of the $z = 3$ UV-bright starburst galaxies are LAEs (e.g., [377, 429, 471]). According to Noll et al. [377], this fraction increases to $\approx 36\%$ for galaxies with $4 < z < 5$. However, this value is based on a relatively small number of objects. For photometrically selected $z \approx 6$ galaxies, LAE fractions between about 20% and 30% have been estimated [144, 493]. This seems to indicate that for $z \geq 3$ the LAE fraction is essentially independent of the redshift. On the other hand, Vanzella et al. [541], who obtained spectra of continuum-selected starburst galaxies in the GOODS-S field, found among a sample of 18 galaxies with spectroscopic redshifts in the range $5.13 < z < 6.20$, 13 objects with conspicuous Lyα emission. According to the publicly available spectra, two of these objects have rest-frame Lyα EWs < 20 Å. Thus, 11 of the 18 objects (or about 60%) seem to be LAEs. An even larger LAE fraction (>80% at $z \approx 6$) has been estimated by Shimasaku et al. [475] from an analysis of LAEs observed in the Subaru Deep Field. The differing results at high redshifts could be due to different selection effects or, possibly, systematic errors. New larger spectroscopic programs, which are in progress, may clarify this question in the near future.

7.1.3.3 Lyα Emission Line Properties

The large majority of the known LAEs have (rest-frame) Lyα emission equivalent widths $\leq$240 Å (e.g., [196, 238]). But there also seems to exist a minority of objects with higher emission equivalent widths. While equivalent widths up to about 240 Å are predicted by the corresponding models of young dust-free starburst galaxies with normal (Salpeter) initial mass functions (e.g., [96, 452]), the objects with the largest EWs must either have very young ages or a different IMF [514].

Most LAEs show Lyα FWHM widths $\leq$500 km s^{-1} (e.g., [124, 514]). None of the reliably measured Lyα line widths exceeds 1000 km s^{-1}. Thus, the Lyα lines of LAEs are much narrower than those of typical high-redshift AGN. Moreover, LAEs normally show no detectable X-ray and radio radiation (e.g., [88, 196]). These facts seem to rule out that ionizing radiation of a central active nucleus contributes to the Lyα flux of the LAEs.

Examples of LAE Lyα line profiles are shown in Fig. 7.37. Basically two types of profiles are observed. Most high-redshift LAEs have Lyα lines with asymmetric profiles, characterized by a steep blue wing and an extended red wing. Profile of this type can be explained by Lyman-forest absorption and/or absorption by galactic winds in the blue wing of an intrinsically symmetric profile. Part of LAEs show double-peak profile, where the two peaks are separated by a narrow and deep apparent absorption feature (e.g., [513, 542]). Models producing such profiles are discussed below in the subsection on the physical nature of the LAEs. There are also LAEs with profiles which appear intermediate between the above two cases, i.e., double-peak profiles with a strong red and a very weak blue peak. A more complete discussion of the observed LAE line profiles and references to the literature on this subject can be found, e.g., in [514].

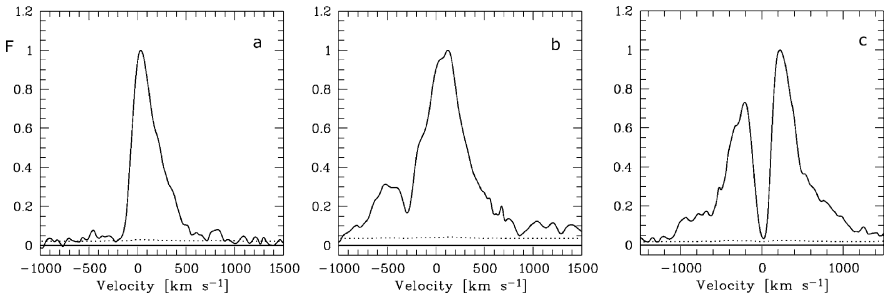

Fig. 7.37 Examples of observed Lyα profiles of LAEs. Most of the LAEs for which spectra of sufficient resolution have been obtained show asymmetric profiles (**a**), with a steep blue wing and an extended red wing. Part of the LAEs show double-peaked profiles (**c**). There are also intermediate cases such the profile (**b**). (Courtesy C. Tapken)

7.1.3.4 LAE Continuum Properties and Sizes

For those LAEs where data of sufficient S/N exist, the rest-frame UV continuum energy distributions of LAEs have been found to be qualitatively similar to those of other UV-bright high-redshift starburst galaxies (e.g., [144]). However, LAEs tend to have steeper FUV continuum slope parameters of $\beta < -2.0$, indicating lower average dust extinction (e.g., [514]). In some cases β is very close to the values expected for completely unreddened starburst.

Compared to other starburst galaxies with similar luminosities and redshifts, the LAEs are systematically smaller (e.g., [144, 514, 544]). Dow-Hygelund and colleagues [144] found for a sample, which included most of the known $z \approx 6$ LAEs

with good HST continuum images, a mean effective radius of 0.60 ± 0.1 kpc. In contrast to the normal high-redshift starburst galaxies, the sizes of the LAEs at $z \approx 6$ seem not to depend on the luminosity.

There is also some evidence that LAEs have on average younger stellar populations than other UV-bright high-redshift starburst galaxies. This may indicate that the Lyα emission decreases with age. A possible cause of this decrease could be dust formation during the course of the starburst activity (e.g., [518]).

7.1.3.5 Stellar Mass and Star-Formation Rates of LAEs

By applying the methods discussed in Sects. 6.3 and 6.6, various authors used the observed continuum properties to estimate masses and star-formation rates of the LAEs. The SFRs derived from the rest-frame UV continuum flux and reported in the literature vary between 1.2 $M_\odot$ yr^{-1} and 50 $M_\odot$ yr^{-1}, with typical values around 10 $M_\odot$ yr^{-1} (e.g., [144, 366, 514]). Thus, LAEs are characterized by modest absolute star-formation rates. As in the case of other star-forming galaxies, the star-formation rates derived from the Lyα flux are generally lower than those estimated from the continuum flux (e.g., [366, 514]). Obviously, in these LAEs part of the intrinsic Lyα flux is internally absorbed. In some LAEs the two methods give similar results. In these objects basically all the intrinsic Lyα flux directed toward us seems to escape.

Less well known are the stellar masses of the LAEs. Mass ranges of $10^6 - 10^8$ $M_\odot$ [420], around $5 \times 10^8 M_\odot$ [196], $2 \times 10^7 - 2 \times 10^9$ $M_\odot$ [171], and $10^9 - 10^{10}$ $M_\odot$ [297] have been reported. Both the largest and the lowest masses have been derived for LAEs at redshifts ≈ 5. Thus, a mass evolution with cosmic time cannot explain the differences. On the other hand, most mass derivations at redshifts $z \geq 5$ are based on fitting SED models to broadband rest-frame UV and visual flux values. As pointed out in Sect. 6.3, such fits can lead to ambiguous results. Moreover they depend on assumptions on the IMF and the star formation histories, which are particularly uncertain for LAEs. A physical explanation for the contradictory results of the LAE mass estimates could be the existence of different types of LAEs. According to Lai and colleagues [298] about 70% of a sample of LAEs observed with the Spitzer satellite at 3.6 μm are young (about 200 Myr) galaxies with stellar masses of the order $3 \times 10^8 M_\odot$. For the remaining galaxies in their sample, Lai et al. find ages $>$ 1 Gyr and masses $\approx 10^{10} M_\odot$.

7.1.3.6 The Chemical Composition of LAEs

Since the LAEs have, on average, fainter continua than other high-redshift galaxies, obtaining spectra of sufficient quality for abundance estimates is difficult. Therefore, at present, there exists no reliable quantitative information on the chemical composition of these objects. However, from the limited spectral information available, it is clear that, contrary to suggestions in the literature, the LAEs are not composed of primordial matter. At least part of the observed LAE spectra show clear evidence for a significant chemical enrichment of the interstellar and stellar matter of these objects. This is demonstrated in Fig. 7.38. In the spectrum reproduced in this figure,

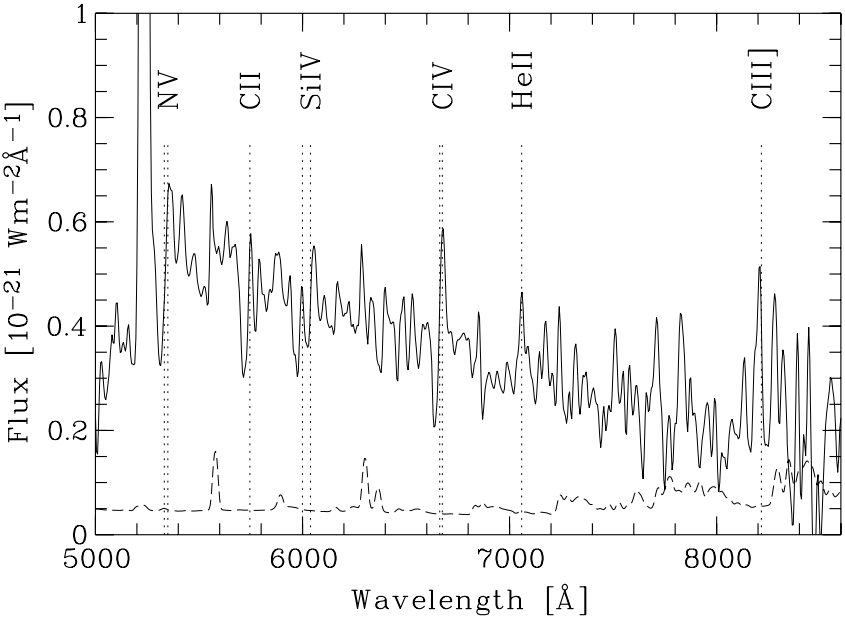

Fig. 7.38 Spectrum of the $z = 3.304$ Lyα emission galaxy of Fig. 7.35 scaled for easier viewing of the absorption spectrum. In addition to hydrogen and helium lines, the C IV and Si IV resonance doublets (showing the typical P Cyg profiles) and interstellar C II absorption features are visible. (From Tapken et al. [513])

the resonance doublets of C IV and Si VI (showing a P Cygni profile typical for very young starbursts) and low-ionization interstellar absorption features are detected. Moreover, although some LAEs have continuum slopes close to the values of dustless galaxies, in other cases the observed continuum can only be explained by the presence of dust absorption. Since the formation of solid dust particles requires chemical elements heavier than H and He, these LAEs cannot be composed of primordial matter. Finally the detailed models of the observed double-peaked Lyα profiles have to assume the presence of some dust.

7.1.3.7 The Physical Nature of the LAEs

According to the observational data discussed above, LAEs are intrinsically faint and small starburst galaxies with modest star-formation rates. The observed high Lyα equivalent widths and a comparison of the SFRs estimated from the UV continuum with those derived from the Lyα flux indicate that in some cases practically all the intrinsic Lyα flux emitted in our direction can escape. But in most LAEs part of the intrinsic line flux is internally absorbed. At least part of the LAEs are composed of matter which has already been enriched with heavy elements. Although the continuum of the LAEs shows, on average, less dust reddening than other high-redshift starburst galaxies, dust is present in at least part of the LAEs.

This finding rises the question why a large fraction of the Lyα photons can escape from these objects, while these photons are internally absorbed in other starburst galaxies. In the literature three basic mechanisms have been suggested which could possibly explain the observations.

The most simple explanation for the high Lyα flux is to assume a geometric structure, where only fully ionized, hot (and consequently dust-free) gas is present in front of the starburst. In this case the Lyα radiation can reach us unattenuated. In view of the small size of the LAEs and in view of the irregular structure of the interstellar gas distribution known from local starburst galaxies (e.g., [291]), this simple model may well apply to some of the observed LAEs.

The explanation given above does not work for the LAEs where the observations indicate that substantial amounts of dust and cool gas are present along the line-of-sight to the continuum and Lyα emitting volumes. Obviously, in these objects the geometric structure and the internal radiative transfer of the UV radiation has to meet special conditions which avoid the quenching of the Lyα radiation by the combination of resonance scattering and dust absorption described at the beginning of this section.

A potential mechanism favoring the escape of Lyα photons has been suggested by Neufeld [373]. This model assumes for the LAEs a two-phase interstellar medium with dense neutral clouds embedded in a thin ionized gas. Furthermore, it is assumed that only radiation which is multiply scattered in the outer layers of the neutral clouds can reach us. Because of the high resonance-scattering cross section, the Lyα photons are (back-) scattered in the surface layers of the neutral clouds, while the continuum photons penetrate deeper into the clouds, where they are absorbed by dust.

The observed small sizes and the low-continuum reddening of most LAEs seem to argue against the Neufeld model. However, for some LAEs observational evidence for a selective escape of the Lyα radiation by the Neufeld mechanism has been reported ([172]). It will be important to clarify the nature of these objects by more detailed spectroscopic observations.

Another mechanism which can avoid the quenching of the Lyα emission is the presence of velocity gradients in the neutral gas. If a hydrogen atom is moving relative to the volume emitting the Lyα line with a velocity larger than the emission line Doppler widths, the atom cannot absorb or scatter the Lyα photons. In this case the Lyα photons can penetrate a neutral hydrogen layer and can escape. An escape of the Lyα photons is also possible if the intrinsic emission line Doppler profile is much broader than the velocity dispersion of the absorbing gas.

Velocity differences between the emitting and the scattering atoms are not unexpected since the velocity of sound is much higher in the ionized Lyα-emitting plasma than in the cool neutral hydrogen gas. Thus, assuming turbulent media with subsonic turbulence, the Lyα line emitted by the hot ionized gas is expected to have a profile which is significantly broader than the velocity dispersion of the cool resonance scattering layers. However, model computations show that in static conditions the absorption in the outer wings of the absorption profiles is sufficient to prevent the escape of Lyα photons. On the other hand, it can be shown that the

probability for an escape of the Lyα photons increases strongly if a large-scale velocity field of the scattering material is present [290]. Mechanisms producing such large-scale velocity fields can be mass infall or mass outflows powered by stellar winds, stellar explosions, or radiation pressure from the central stars or star clusters (e.g., [123, 339, 515]).

Since velocity fields affect the profiles of the Lyα lines, line profile observations provide a critical test of the hypothesis that large-scale velocity fields are responsible for the escape of the Lyα photons. As shown by several different authors, the observed Lyα profiles and the emission strengths of LAEs can indeed be reproduced well by models of this type (e.g., [6–8, 348, 514, 545, 546]). An example of such model comparisons is given in Fig. 7.39. The corresponding model assumes a central volume of turbulent, ionized hydrogen, emitting the Lyα radiation. This region is assumed to be surrounded by a shell of dusty neutral hydrogen. To reproduce the observed profile a modest (≈ 10 km s^{-1}) expansion velocity of the neutral hydrogen shell had to be assumed. From the observed line wings (which are not much affected by the neutral hydrogen shell) supersonic turbulent velocities (> 600 km s^{-1}) of the ionized gas are estimated. The high velocities are assumed to result from supernova shocks and strong stellar winds in compact, strong starbursts. Because of the relatively broad intrinsic Lyα profile and the expansion of the neutral hydrogen shell, a major fraction of the intrinsic Lyα photons can escape in spite of the presence of a dusty neutral hydrogen shell in front of the emitting region.

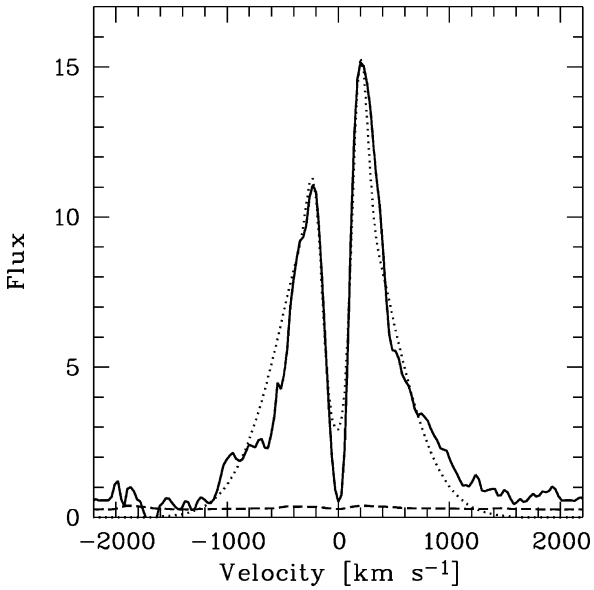

Fig. 7.39 Comparison of an observed Lyα line profile of the LAE FDF-4691 (*solid line*) and a profile predicted by a model assuming a central turbulent ionized emission region surrounded by an expanding neutral envelope (*broken line*). (From Tapken et al. [513])

7.1.3.8 Extended Lyα Sources ("LABs")

During the searches for distant Lyα emitters occasionally large, spatially extended line emission regions are observed. The first of these objects were found in 2000 by C. Steidel and colleagues [497] in a region of unusually high galaxy density at a redshift of $z = 3.09$. Following Steidel et al. these objects are called "Lyα blobs" (LABs). Additional LABs were found by chance or during special search programs devoted to these objects (e.g., [130, 341, 375, 443]). An example of an LAB is presented in Fig. 7.40. According to the observations, LABs in the redshift range $2.7 < z < 5$ have sizes between ≈ 10 kpc and ≈ 200 kpc. Their Lyα luminosities range between about 10^{35} W and about 10^{37} W. Thus, LABs are not only much larger than the LAEs, but they are also more luminous. In contrast to the LAEs, normally no continuum emission can be detected in LABs. Some LABs appear to be associated with high-redshift galaxies or AGN (e.g., [198]); others seem to be completely isolated. In one case, in addition to Lyα, extended line emission of C IV and He II has been reported [130].

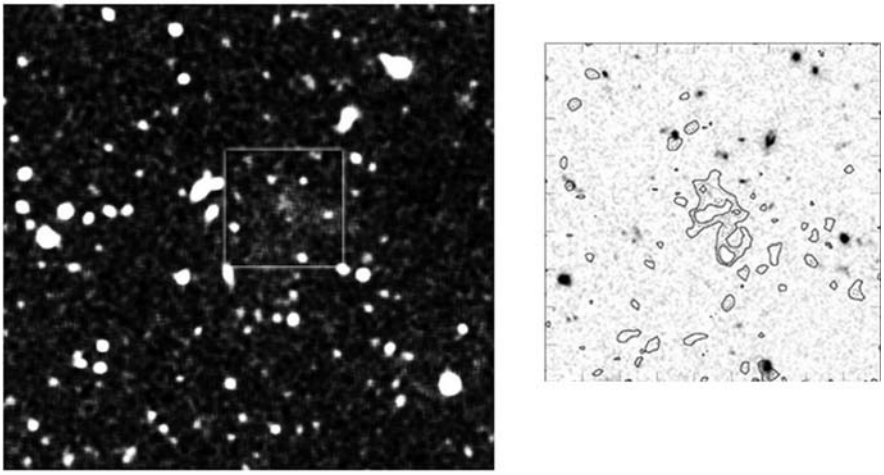

Fig. 7.40 A Lyα blob at $z = 3.16$. The *left image* is a VLT FORS1 narrow-band image centered on 5055 Å. The *right image* shows a contour map of the redshifted Lyα emission in the central region of the *left image*, superposed on an HST image of the field. The size of the *right image* is 18"×18" (corresponding to about 133 kpc × 133 kpc at this distance). For details on this LAB see [375]. (Image: European Southern Observatory, Press Image 23/06)

Because of the relatively small number of well-studied LABs and because of the limited information available, the physical nature of the LABs is not at all clear at present. Among the suggested explanations put forward in the literature is the proposal that LABs are extended nebulae, ionized by bright galaxies or by (for us) dust-obscured AGN. Other authors suggested that in the case of the LABs we observe shock-heated gas in regions of violent galaxy formation or mergers. Another possible explanation is that we are observing cooling radiation from matter

being accreted by a massive dark-matter halo. More information is needed to clarify the nature of these rare but interesting objects, and to determine the relation between the LABs and the high-redshift galaxies.

7.1.4 Dust-Obscured Starburst Galaxies

As noted in the preceding section, the radiation of many high-redshift starburst galaxies is attenuated and reddened by internal interstellar dust absorption. The observed amount of dust reddening varies greatly from galaxy to galaxy. While some starburst galaxies (in particular part of the LAEs) show no detectable reddening, there exist high-redshift galaxies where the dust absorption fully obscures the UV and optical light. From these galaxies only redshifted FIR and radio radiation is observed. These dust-obscured galaxies have initially been identified by their (observer-frame) sub-millimeter or millimeter wavelength flux. Therefore, they are normally referred to as "submm galaxies" or SMGs (Sect. 2.2.6). Spectroscopic redshifts of such "optical dropouts" can be obtained from molecular lines emitted at infrared, submm, or radio wavelengths. Photometric redshifts can be derived from the characteristic energy distribution of the FIR emission of the dust, or from distinct dust emission features (e.g., [244]). However, the photometric redshifts based on the dust emission spectrum are affected by the dust temperature (which normally cannot be determined independently). Therefore, these photometric redshifts are less accurate than photometric redshifts based on optical data.

The present detectors for submm imaging observations have small fields of view. Therefore, blind searches for submm sources are inefficient. Instead such objects are normally discovered by observing positions where obscured starburst galaxies are expected because of the presence of strong radio emission with properties characteristic of massive young starburst. Thus all major samples of distant submm galaxies consist of radio-preselected objects.

Not all submm-bright high-redshift galaxies are "optical dropouts." Some of the galaxies initially identified from their high submm flux can also be detected using the optical or infrared search techniques. According to Reddy and colleagues [427], more than 80% of the radio-selected submillimeter galaxies at $z \approx 2$ are also found using the Lyman-break, the BzK, or other NIR search methods for identifying distant starburst galaxies.

Many submm-bright high-redshift dusty starburst galaxies, for which the attenuated UV and visual spectrum could be detected, show (apart from the reddening) spectroscopic properties similar to those found in the UV-bright systems. Normally, similar star-formation rates are deduced from the dereddened UV flux and from the FIR dust emission (e.g., [119, 507]). However, in contrast to the (rest-frame) UV and optically selected galaxies, the submm-bright objects often show evidence for an AGN contribution to the spectrum. But the presence of polycyclic aromatic hydrocarbon (PAH) features in the IR spectra of high-redshift submm sources (Fig. 7.41) indicates that starburst activity contributes significantly to the IR flux (e.g., [324, 463]). In the most luminous dust-obscured starburst galaxies, the

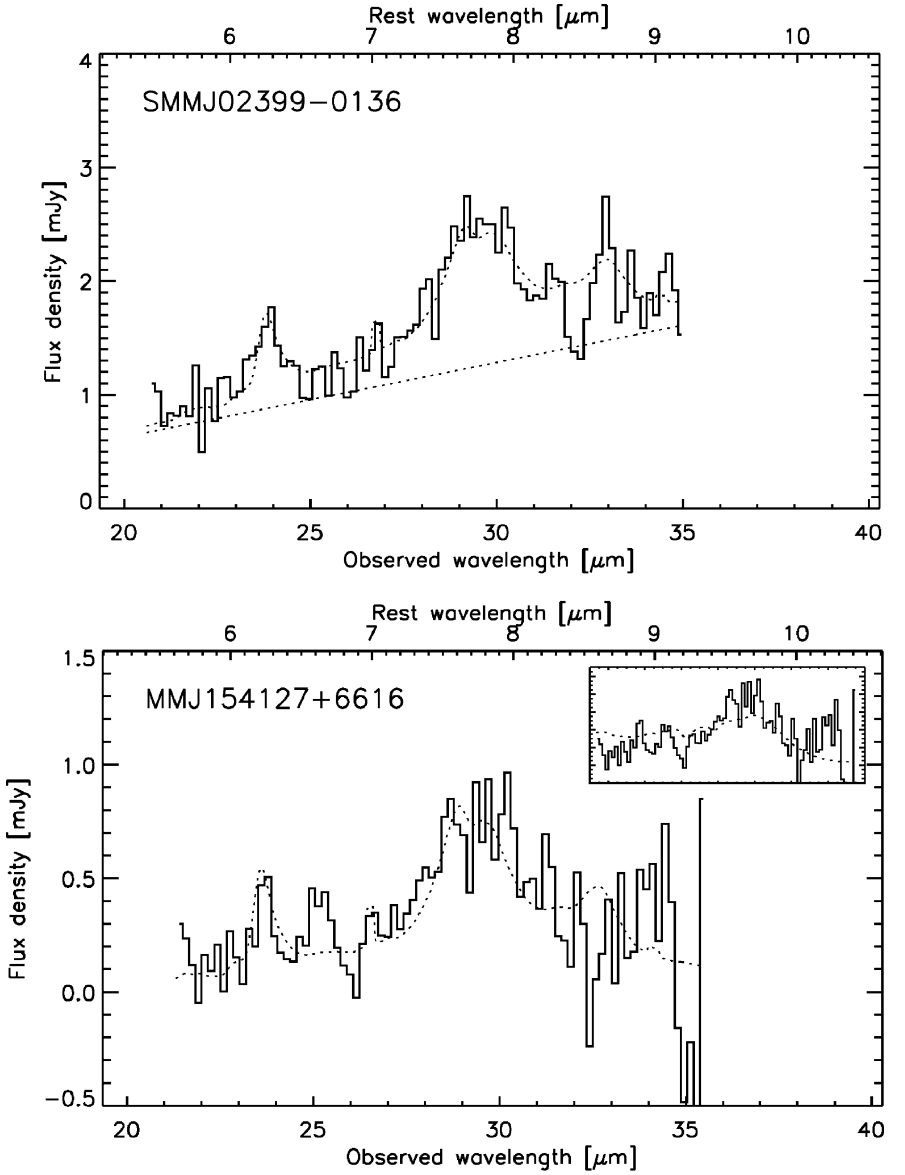

Fig. 7.41 The 6.2 and 7.7 μm polycyclic aromatic hydrocarbon (PAH) emission bands in the spectra of two $z = 2.8$ submm galaxies with AGN cores (*solid line*). For comparison the (scaled) spectrum of the local starburst galaxy M 82 is overplotted (*dotted spectrum*). For the comparison with SMM J02399-0136, an AGN continuum has been added to the M 82 spectrum. (From Lutz et al. [324])

star-formation rates derived from the dereddened UV flux tend to be significantly lower than those derived from the FIR flux. The reasons for this discrepancy could be special dust properties or a complete obscuration of the regions of maximal star formation at UV wavelengths. Obviously, in these cases the dereddened UV flux results in an underestimate of the SFRs.

Those high-redshift submm galaxies, for which reliable redshifts could be determined, are starburst galaxies with high star formation rates of the order $10^2 M_\odot$ yr^{-1} to $>10^3$ $M_\odot$ yr^{-1}. In general the dust-obscured starburst galaxies appear more massive and more luminous than the average high-redshift starburst galaxies.

At present, significant samples of dust-obscured starburst galaxies are known at redshifts up to about $z = 3$ [246]. Their typical redshift is about $z \approx 2$. For higher redshifts, dust obscuration in galaxies appears to decrease rapidly with redshift. This decrease probably results from the lower abundance of heavy elements (which are a prerequisite for dust formation) in the higher redshift objects. But there do exist individual highly dusty submm galaxies at redshifts $z > 4$ (e.g., [87, 120]). Of particular interest is a $z = 4.547$ submm galaxy with an estimated stellar mass $> 10^{10} M_\odot$ and a SFR $> 10^3 M_\odot$ yr^{-1} [87].

Those high-redshift submm galaxies, for which high spatial resolution observations are available, were found to be very compact, showing higher densities of the starburst regions than the UV-bright high-redshift galaxies. At least some of these objects fit the criteria of "maximal starbursts," in which practically all the available initial interstellar gas in converted into stars on (or close to) the local dynamical timescale. After the exhaustion of the gas the intense star formation has to cease. This means that these luminous submm galaxies are shortlived ($\leq 10^8$ years) events. The observed CO gas kinematics suggest that we observe in all these cases highly dissipative mergers of gas rich galaxies [505, 506].

7.2 Passively Evolving High-Redshift Galaxies

Because of the absence of young hot stars, passively evolving galaxies show only little rest-frame UV radiation. As a result, these galaxies were not detected in the early high-redshift galaxy surveys, which were designed to find UV-bright objects. Because of the drop of the spectral flux shortward of (rest-frame) 4000 Å (Fig. 6.3), passively evolving galaxies are difficult to observe with silicon-based CCD detectors if they are at redshifts $z > 1.5$. But they are readily found in galaxy samples targeted at very red and NIR-bright objects.

Samples of distant galaxies with red colors generally include passively evolving objects as well as dust-reddened star-forming systems. Samples selected by their NIR color index $J - K > 2.3$ (called "distant red galaxies" or "DRGs;" [178]) are known to consist mostly of dusty star-forming galaxies. But observations with the Spitzer infrared satellite have shown that about 30% of the DRGs have rest-frame NIR colors consistent with passively evolving galaxies at redshifts $2 < z < 3$ [295]. More recent studies of the DRGs give, within the error limits, similar results [282, 427].

With suitably selected filter bands, NIR photometry can be used to directly separate the passively evolving galaxies from reddened star-forming objects. An example is the BzK selection technique described in Sect. 5.1.2.

Galaxies selected using the "extremely red object" ("ERO") criteria ($R - K > 5$ or $I - K > 4$) normally have redshifts <2 (e.g., [118, 141]). Like the DRGs, the EROs include passively evolving galaxies as well as reddened star-forming galaxies.

7.2.1 Properties of the High-Redshift Passively Evolving Systems

Passively evolving high-redshift galaxies were initially identified on the basis of SED fits to synthetic spectra. These identifications were confirmed by NIR spectra of these objects (e.g., [281, 282, 284]). As noted by M. Kriek and colleagues [281] (and as shown in Fig. 7.42), galaxies identified from multiband photometry as passive systems, as a rule do not show detectable H II region emission lines. This seems to confirm that no hot young stars are present in these galaxies. A more detailed analysis shows that for part of the identified passively evolving objects the

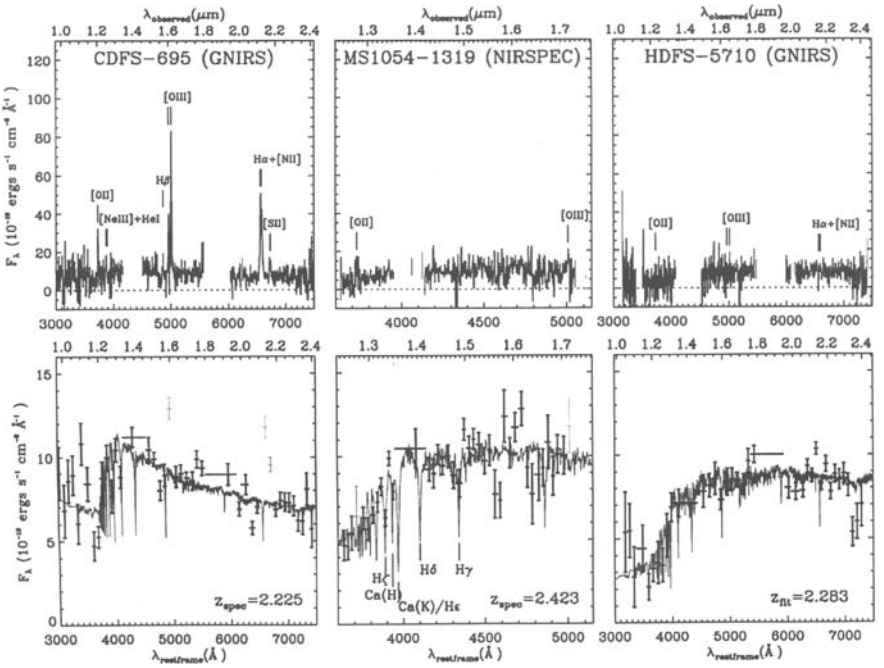

Fig. 7.42 Examples of (rest-frame) optical spectra of $z \approx 2.3$ NIR-selected galaxies. The *upper panels* give observed spectra. The data points in the *lower panels* give the observed flux averaged over 25 or 13 Å bins (in the observers frame). Overplotted in the *lower panels* are best-fit template spectra. As shown by the blue SED and the H II region line emission, CDFS-695 is a star-forming galaxy, while HDFS-5710 shows all spectral properties of a passively evolving system. MS 1054-1319 is an intermediate case. (From Kriek et al. [281])

energy distributions agree well with the assumption of negligible star-formation activity. But there is also a substantial fraction of galaxies which (in addition to an old population) seem to contain some young and intermediate-age stars. Such galaxies are also observed at lower redshifts and are referred to as "E+A" galaxies (e.g., [141]). The presence of some low-level star formation in part of the apparently passively evolving high-redshift galaxies may also be indicated by warm dust emission (typical for star formation regions) from some of these objects. According to P. Perez-Gonzalez and colleagues [410] at $1 < z < 2$, a majority of the massive galaxies with colors typical for passively evolving systems show MIR warm dust emission. On the other hand, A. Cimatti and colleagues [99], studying a different sample of (according to the photometric data) passively evolving galaxies in the same redshift range found warm dust emission almost always absent in such systems.

More accurate constraints on the population content of the passively evolving high-redshift galaxies can be obtained from their line spectra and from a detailed comparison of the continua with model spectra. An example of such a comparison is given in Fig. 7.43. The figure shows a low-resolution ($R \approx 600$) composite spectrum of 13 galaxies in the redshift range $1.4 \leq z \leq 2.0$ with IR colors corresponding to a passively evolving stellar population. The figure shows that the observed spectrum

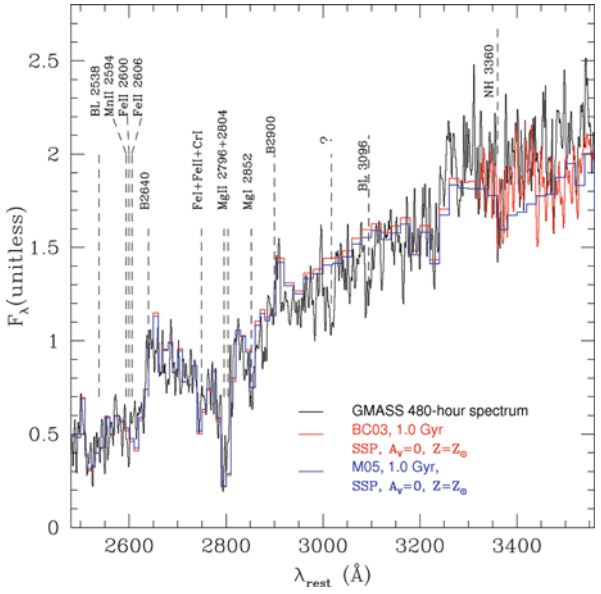

Fig. 7.43 Composite (rest-frame) UV spectrum of 13 passively evolving galaxies with redshifts $1.4 \leq z \leq 2.0$ ($\bar{z} = 1.634$). Also included in the figure are synthetic spectra of a 1 Gyr old passively evolving stellar population, as derived by Bruzual and Charlot ([78], BC03) and by C. Maraston ([336], M05) for a solar chemical composition. As shown by the figure, the observed composite spectrum is fitted reasonably well by these models. (From Cimatti et al. [99])

agrees rather well with that predicted for a passively evolving stellar population with an age of about 1 Gyr.

As shown by Kriek et al. [284], the passively evolving galaxies with a mean redshift $z = 2.3$ already form a "red sequence" (Sect. 2.2.4) in the (rest-frame) $U - B$ vs. stellar-mass diagram. Compared to the corresponding diagram for elliptical galaxies of the local universe, the red sequence at $z = 2.3$ is shifted to smaller values of $U - B$ by about .15 magnitudes. Moreover, as illustrated by Fig. 7.44, the red-sequence galaxies at $z = 2.3$ show spectroscopic evidence for a significantly younger stellar age. This seems to indicate that at the epoch corresponding to $z = 2.3$ the red sequence had just been starting to build up.

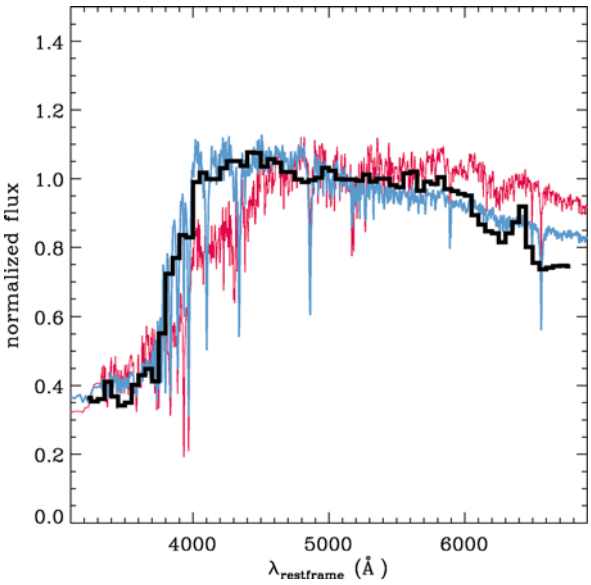

Fig. 7.44 Composite spectrum of $z \approx 2.3$ red-sequence galaxies (*thick black line*). Overplotted is a mean of the template spectra which provided the best fit of the individual observed galaxy SEDs (*blue line*). Also included is a synthetic spectrum assuming an about 2 Gyr old passive population (*red line*). The synthetic spectrum has less flux in the region $4000 < \lambda < 4500$ Å, and more flux at $\lambda > 5500$ Å than the observed spectrum. The comparison indicates that the observed population is younger than 2 Gyr. (From Kriek et al. [284])

A formation epoch of the red sequence corresponding to the redshift range $2 < z < 3$ is also supported by Fig. 7.45, which shows the rest-frame color index $U - V$ as a function of the absolute blue luminosity for an IR-selected sample of galaxies with redshifts ≤ 3.5. These galaxies were investigated by S. Salimbeni and colleagues [444], using data from the GOODS-MUSIC catalog. As shown by Fig. 7.45, the bimodal distribution of the galaxy colors is clearly visible for $z \leq 2$, and some bimodality is already indicated for the $2 < z < 3.5$ galaxies. The figure also shows that the blue *and* the red galaxies become on average somewhat bluer with increasing redshift.

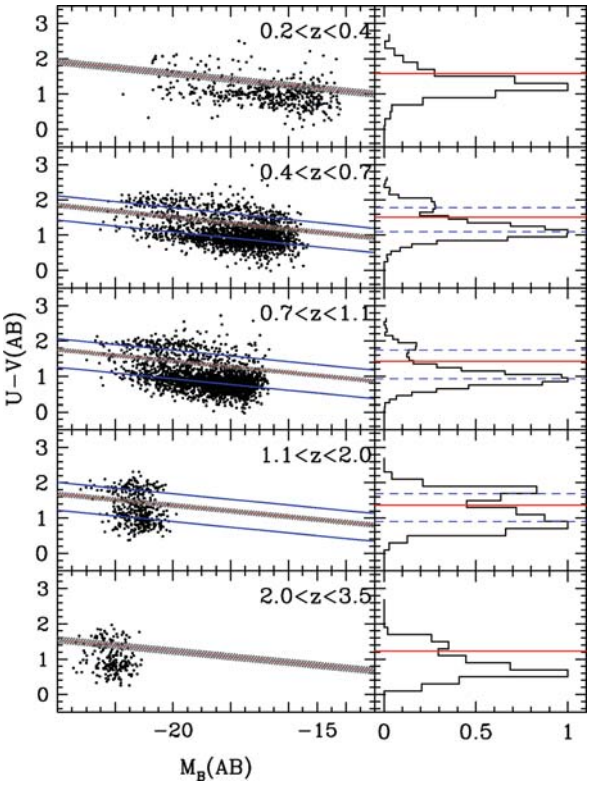

Fig. 7.45 The rest-frame color index $U - V$ as a function of the absolute blue magnitude M_B in various redshift intervals for a NIR-selected galaxy sample (*left panel*). The *thick shaded line* in the *left panel* indicates the minimum in the data point density which separates the red sequence from the blue galaxy population. The *right panel* gives histograms of the color distributions projected at $M_B = -20$. The *solid lines* in the *right panel* separate the blue and red galaxies at this luminosity, while the *dashed lines* indicate the maxima of the distributions. (From Salimbeni et al. [444])

When interpreting Fig. 7.45 selection effects have to be taken into account. At high redshifts the sample is truncated by the magnitude limit of the observations. But this is not expected to influence the conclusions concerning the bimodal color distribution of the galaxies (even) at high redshift.

As described in the paper by Salimbeni et al. [444], the sample of Fig. 7.45 also shows at all redshifts a well-pronounced bimodality in the distribution of the star-formation rates [444].

As pointed out in Sect. 2.2.2, in the local universe passively evolving systems dominate the galaxy number density at high mass. Interestingly, such a trend is observed already at high redshifts. Most of the known high-redshift passively evolving systems have stellar masses $> 10^{11} M_{\odot}$. Moreover, passively evolving galaxies are found most frequently in galaxy samples which are selected to have high stellar mass values (e.g., [111]). The tendency of the passively evolving galaxies to be massive seems to be in place at least by the epoch corresponding to $z = 2$.

A notable property of the high-redshift passively evolving systems is their often very high mass density (e.g., [99, 526, 538]). The observed stellar mass densities are higher by factors of the order 10^2 than the densities of present-day elliptical galaxies. During a passive evolution without mergers the mass density cannot decrease. Thus, observed dense high-redshift systems cannot have passively evolved to become the typical lower density elliptical galaxies observed in the local universe.

7.2.2 The Space Density of Passively Evolving Galaxies

The samples of well-established passively evolving galaxies with redshifts $z > 2$ are still small and not always well defined. Therefore, their space density is still somewhat uncertain. Nevertheless, it is clear that at $z > 2$ the number density of the bright passively evolving systems has been much lower than in the present-day universe (e.g., [295, 504]). According to Labeé and colleagues [295] at present the (comoving) number density of passively evolving systems is about 10 times higher than at the epoch corresponding to $z = 2.5$. This again confirms that (at least most of) the local elliptical galaxies cannot have resulted from a simple passive evolution of the passively evolving galaxies observed at high redshift. A similar conclusion is derived from the properties of the red-sequence galaxies observed at $z = 2.3$ [284].

7.3 Quasar Host Galaxies and Radio Galaxies

As pointed out in Chap. I, with present techniques luminous quasistellar objects can be observed at any distance and any redshift at which they are expected to exist. Moreover, because of their distinct spectral properties, QSOs can be easily and reliably identified. Thus, in principle, observing the host galaxies of QSOs provides a powerful tool to investigate galaxy properties at high redshift. On the other hand, per definition, QSOs are systems where the flux received of the stars and of the interstellar matter is much smaller than that of the central active nucleus. Analyzing the properties of the corresponding galaxies requires the subtraction of the strong AGN contribution from the observed spectral flux. With linear detectors this is possible, but the subtraction introduces statistical (and sometimes systematic) errors into such measurements. As a result, these observations are difficult and time consuming.

High angular resolution studies of local QSO host-galaxies show that QSOs can occur in all types of galaxies [26, 67]. However, in the local universe the most luminous and the radio-loud QSOs usually reside in massive elliptical galaxies.

Examples of images of low-redshift QSO host galaxies are reproduced in Fig. 7.46. As shown by this figure, most images of local host galaxies of QSOs show evidence of ongoing or recent mergers or of galaxy interactions. As has been

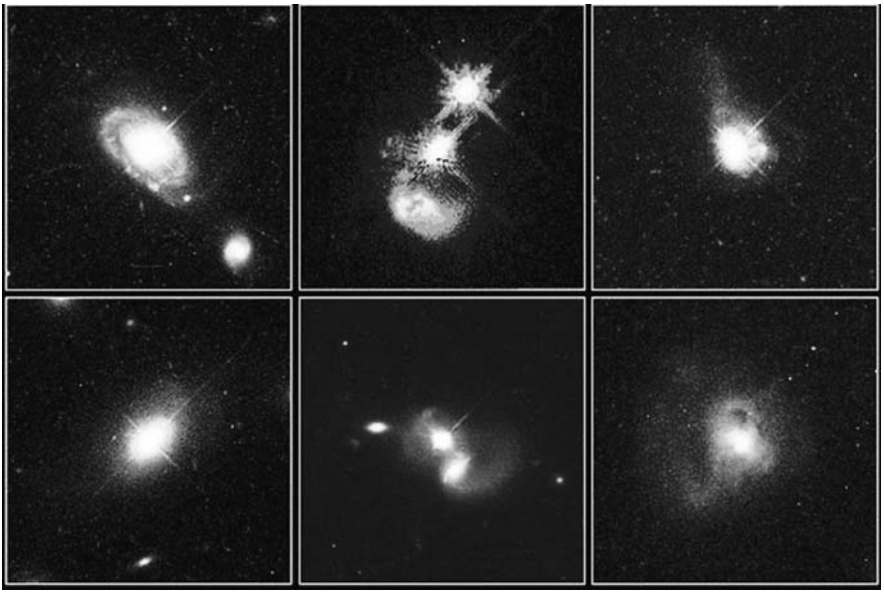

Fig. 7.46 HST images of local QSO host galaxies. The two galaxies on the left are a relatively normal disk galaxy and an elliptical galaxy. The morphology of the other four host galaxies indicate galaxy mergers. (Credit: John Bahcall (Institute for Advanced Study, Princeton), Mike Disney (University of Wales), and NASA)

pointed out in Sect. 2.2.7, an enhanced mass accretion of the central black hole resulting from the mergers probably causes the QSO activity.

Optical and NIR observations of host galaxies with redshifts $2 < z < 3$ have been made with ground-based instruments as well as with the HST. In most cases the observed galaxy properties were found to correspond to those of luminous (1 to 10 L^*) and massive passively evolving galaxies (e.g., [164, 278, 283, 289, 434, 461, 548]). However, compared to the massive elliptical galaxies which host the local bright QSOs, the host galaxies of $z \geq 2$ QSOs are less massive and have higher central densities. Like in the local universe, the host galaxies of radio-loud QSOs are on average brighter and more massive than the hosts of radio-quiet QSOs. Rest-frame UV observations of the high-redshift QSO hosts indicate that these galaxies typically contain (in addition to the old stars) a significant fraction of newly formed stars [249, 461]. The estimated star-formation rates (of the order 6 $M_\odot$ yr^{-1} are comparable to those of the (less massive) Lyman-break galaxies.

Sub-millimeter and millimeter-wave observations of $z \approx 6$ QSOs prove the presence of a surprisingly large amount of dust and cool, molecular gas in these objects (e.g., [50, 89, 90, 331, 550, 552]). The luminosities as well as the BLR line profiles of these distant QSOs indicate black hole masses up to $5 \times 10^9 M_\odot$ (e.g., [292]). Molecular gas masses $>10^{10} M_\odot$ and dust masses $\geq 10^8 M_\odot$ have been estimated for these objects. These numbers provide firm lower limits of the baryonic mass of these high-redshift QSO host galaxies. According to numerical simulations, $z > 6$ QSOs

with the observed properties and space densities require dark matter halos of about $6 \times 10^{12} M_{\odot}$.

The molecular gas mass of the high-redshift QSO hosts has been determined from the line emission of CO molecules. Hence, these observations also show that the interstellar medium of these QSO host galaxies has been significantly enriched with oxygen and carbon (and presumably other heavy elements). Moreover, the formation of dust is possible only if elements heavier than helium are present. Thus, a substantial chemical enrichment of the ISM of the distant QSO host galaxies must already have taken place during the first Gyr of cosmic time. To produce the observed large mass of heavy elements, star formation must have begun rather early in these massive QSO host galaxies.

The early chemical enrichment of the interstellar medium of the high-redshift host galaxies of QSOs is also evident from the chemical composition of the BLR gas of the high-redshift QSOs (e.g. [218]). Even at the highest redshift approximately solar or super-solar chemical abundances of the BLR gas are observed. Furthermore, observations of the Fe/Mg ratio show that the BLR gas of these objects contains a normal ratio between iron and α-elements (e.g., [136, 292]). This indicates an element synthesis similar to that in the Milky Way galaxy.

In the Milky Way and in local star-forming galaxies, dust grains are produced most efficiently in the cool atmospheres of evolved low-mass stars in the "asymptotic giant branch" (AGB) evolutionary phase. To reach this stage a stellar-evolution age of at least 1 Gyr is needed. Since a redshift of $z = 6$ corresponds to an cosmic age of about 0.9 Gyr only, dust formation by AGB stars can be ruled out for the $z > 6$ QSO host galaxies. In these objects all the observed dust has probably been formed in SNe. The different formation mechanism suggests that the dust properties in these distant objects may be different. This seems to be confirmed by infrared observations of the QSO SDSS J104845.05+463718.3 ($z = 6.193$), which shows a dust extinction law that is quite different from that of local and low-redshift objects [332].

From the observed SEDs of the $z > 5$ QSO host galaxies, dust temperatures of the order 40–50 K have been estimated. Since the AGN radiation field may contribute to the heating of the dust, SFRs cannot be determined unambiguously from the dust emission. However, the observed FIR spectral energy distribution of these objects is typical for star-forming galaxies [552]. A significant starburst contribution to the light from these objects is also indicated by the presence of polycyclic aromatic hydrocarbon (PAH) emission features (Fig. 7.41) in the IR spectra of the high-redshift QSO host galaxies (e.g., [323]). For the host galaxy of the QSO SDSS J114816.64+525150.3 ($z = 6.42$) Maiolino and colleagues [329] derived from measurements of the [C II] 158 μm fine structure line a star-formation rate of $\approx 3000 M_{\odot}$ yr^{-1}. Thus, the host galaxies of the very distant QSOs not only harbor very massive and luminous AGN, but they also experience particularly intense star formation activity.

According to the unified model of the AGN, radio-loud QSOs and radio galaxies differ only by the orientation of their symmetry axis relative to the line of sight. Therefore, distant radio galaxies are expected to have similar properties and stellar

populations as the host galaxies of the radio-loud QSOs. This is basically confirmed by the observations. Infrared observations with Spitzer of a sample of 69 radio galaxies with redshifts $1 < z < 5.2$ show that in most high-redshift radio galaxies the observed rest-frame IR flux is dominated by stellar light. The stellar masses of these galaxies were found to range approximately between $10^{11} M_{\odot}$ and $3 \times 10^{11} M_{\odot}$ [469]. At least in one high-redshift radio galaxy PAH emission has been detected [468].

7.4 The Host Galaxies of Distant Gamma-Ray Sources

The long-duration Gamma-ray bursters (Sect. 2.1.5) are the most intense radiation sources in the universe. Therefore, like QSOs, they can be observed over very large distances. Moreover, due to their unique properties they can be readily distinguished from foreground objects. Finally, since from observations it seems to be established that long GRBs are found preferably in low-metallicity environments, they are expected to occur more frequently in the less chemically enriched high-redshift galaxies.

Like in the case of the QSOs, the spectral energy distribution of GRBs covers a wide wavelength range. Because of the extreme brightness of the bursts, the light of their host galaxies cannot be directly observed during the outburst (see, e.g., Fig. 7.47). However, since the Gamma-ray bursts are transient phenomena, in principle, the host galaxies can be studied as soon as the light curve of the burst has sufficiently decayed. Moreover, the interstellar medium of the host galaxies in front

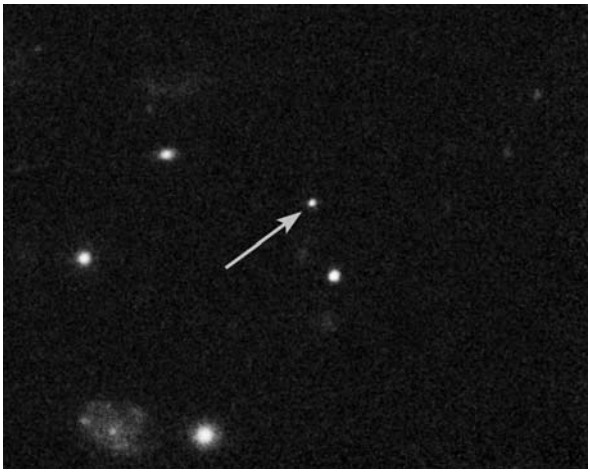

Fig. 7.47 The 16 arcsec × 12 arcsec section of an HST image of the Gamma-ray burst source GRB 080319B (indicated by the *white arrow*). Although the HST image has been taken almost 3 weeks after the outburst, the afterglow of the burst was still so bright that no trace of the underlying galaxy could be detected in this image. (Credit: NASA, ESA, N. Tanvir (University of Leicester), and A. Fruchter (STScI))

of the outburst source can be investigated using interstellar absorption lines and dust absorption features in the spectrum of the outburst or its afterglow. Such observations are facilitated by the very high radiation flux observed during the outbursts. On the other hand, the short duration of the outbursts require a rapid response, and the fast decay of the burst light curves limits the integration time. The short timescale of the outbursts is less critical in the case of high-redshift GRBs, since the outburst durations become longer $\propto (1 + z)$ in the observer's frame.

High-resolution HST observations show that long-duration GRBs often occur in low-mass star-forming galaxies with low chemical abundances [183]. In the low-redshift universe the typical hosts of long GRBs are small, low-metallicity irregular galaxies. There is some disagreement in the literature on how closely GRBs are associated with concentrations of young stars. Fruchter and colleagues [183] find the GRBs strongly concentrated on the very brightest regions of their host galaxies. Hammer et al. [219] found from high spatial resolution images of a sample of GRB hosts that the outbursts typically occur in the range of 400–800 pc from the regions of highest stellar density.

Keck spectroscopy of the optical afterglow of the GRB 050505 ($z = 4.275$) resulted in a metallicity of the interstellar medium of the GRB host galaxy of $Z \approx 0.06 Z_\odot$ [48]. This is well below the value expected from Fig. 7.21 for high-redshift starburst galaxies at this redshift. It seems to confirm that GRB host galaxies have a lower than average metallicity, even at high redshift. The relative abundances of the observed elements in the host galaxy of GRB 050505 were found to show a dust depletion pattern similar to that observed in the diffuse interstellar gas of the Milky Way. Evidence for dust depletion has also been reported by J. Prochaska and colleagues, who studied the inrinsic ISM absorption lines in the spectra of 16 GRBs with redshifts $1.255 < z < 6.296$ [422]. However, in their sample Prochaska et al. found metallicities which appear more normal for the observed redshifts. While all objects showed subsolar metallicities (between about 0.7 and 0.005 of the solar value), on average no significant deviation from the values expected at the corresponding redshifts could be detected.

All spectroscopic observations of GRB afterglows show a high H I column density toward the GRBs. This confirms that long-duration GRBs occur in very dense environments.

Using photometry over a large wavelength range, Stratta and colleagues have investigated the dust properties of the $z = 6.29$ host galaxy of the GRB 050904 [503]. They found that the observed dust reddening cannot be explained by the dust extinction law derived for local or lower-redshift galaxies. On the other hand, the observed dust reddening appears to be compatible with the extinction law derived earlier for the $z = 6.193$ QSO SDSS J104845.05+463718.3 (Sect. 7.3). This seems to confirm that for galaxies with redshifts $z > 6$ different dust properties need to be assumed.

Theoretical models for the formation and evolution of the GRB host galaxies have been computed by A. Lapi and colleagues [300]. For their simulations these authors assumed that in galaxies with $Z < 0.2 Z_\odot$ about 2% of the stars with masses in the range $8 M_\odot < M < 100 M_\odot$ end their life as GRBs. With these assumptions

the observed statistics of the long-duration GRB could be well reproduced. According to the simulations, long GRBs occur predominantly in star-forming galaxies with a low halo mass ($\leq 10^{11} M_{\odot}$) and a young (<50 Myr) stellar population. Because of the metallicity dependence of the occurrence of GRBs, the cosmic rate of GRBs is predicted to reach a maximum at higher redshifts (at $z \approx 6$) than the cosmic star-formation rate. About 30% of all long GRBs are predicted to occur before $z = 6$, and about 60% are expected to take place before $z = 4$. These results are consistent with the assumption that the long-GRB host galaxies form the low-mass and low-metallicity end of the UV-bright, high-redshift star-burst galaxies. However, in view of the somewhat contradictory observational results on the average metallicity of GRB host galaxies mentioned above, additional observations will be needed to confirm this assumption.

Chapter 8
The Space Distribution of High-Redshift Galaxies

As discussed in Chap. 3, the matter distribution, which we observe in the present-day universe in the form of galaxies and intergalactic gas, is the result of the formation and evolution of structure at early cosmic epochs. This evolution is still in progress. The currently accepted CDM model of the evolution of the cosmic structure makes detailed predictions of the dark and baryonic matter distribution as a function of cosmic time and redshift (Figs. 3.4 and 3.5). Galaxies are expected to trace this distribution. Thus, observations of their space distribution as a function of redshift can provide valuable cosmological information.

8.1 The Galaxy Number Density as a Function of Distance

For the local universe the sponge-like large-scale structure of the galaxy distribution has been known since several decades [466]. Evidence for similar large-scale structure of the distribution of the high-redshift galaxies was first detected in the redshift distribution of galaxies in deep-field surveys (e.g., [103]). Because of the redshift-distance relation, the redshift distribution of galaxies in a certain direction traces the space distribution of these objects along the corresponding line of sight. Thus, galaxy redshift distributions allow us to study the space distribution of galaxies in one dimension. If the galaxy distribution follows the predicted density structure, we expect density peaks where the line of sight intersects a wall or filament of the matter distribution, and regions of low galaxy number density when the LOS passes through a void. An example of observed redshift distribution at $z > 2$ is given in Fig. 8.1. As expected from the theory, we observe statistically significant overdense regions (in Fig. 8.1 at $z = 2.35$, 3.13, and 3.4) and empty regions (e.g., at $2.8 < z < 3.05$). Within the statistical uncertainties, the observed scales are in agreement with the theoretical predictions (e.g., Fig. 3.4). A more surprising result is the narrowness of some of the density peaks.

Galaxy redshift distributions of the kind shown in Fig. 8.1 have been observed in many different directions. Examples for different fields with $2 < z < 3.5$ can be found in [495] and [498]. An example of an galaxy overdensity visible in a redshift distribution diagram at $z > 5$ is reproduced in [334].

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_8,

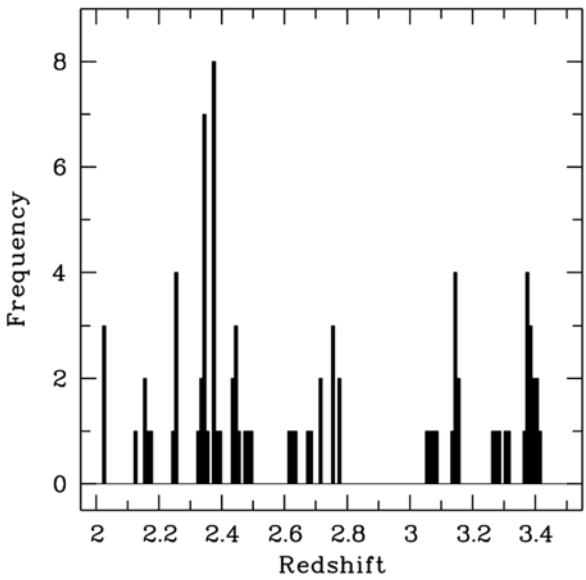

Fig. 8.1 Distribution of the galaxy number density as a function of redshift for galaxies with $2 < z < 3.5$ in the FDF spectroscopic sample. The bin width is $\Delta z = 0.01$. (From Noll et al. [377])

The analysis of the redshift distribution allows us to study the number density structure of galaxies over large distance intervals. But a disadvantage of this technique is that only one dimension is covered. Therefore, the interpretation of the data has to rely on model assumptions. Moreover, statistically representative results from redshift distributions require galaxy samples with accurate spectroscopic redshifts.

8.2 Statistics of Galaxy Positions

Most of the available information on the space distribution of high-redshift galaxies has been derived by analyzing the distribution of the positions of these objects on the sky. Since a large majority of the high-redshift galaxies have been identified from deep-field searches, which typically cover small areas, most of the published results describe the small-scale properties of the galaxy distribution.

A convenient measure of the clustering properties of the galaxies in two dimensions is the "angular two-point correlation function" $w(\theta)$, which is defined by

$$N(\theta) = N_G(1 + w(\theta)) \tag{8.1}$$

where $N(\theta)$ is the number of galaxies observed on the sky per unit solid angle as a function of the angular distance θ from any given galaxy and N_G is the mean surface number density. $w(\theta)$ describes the excess probability of finding a galaxy at an angular distance θ from any given galaxy. In the case of a completely uniform random distribution of the galaxies, we have $w(\theta) \equiv 0$ and $N(\theta) \equiv N_G$.

From observational data, $w(\theta)$ can be calculated directly by means of the definition Eq. (8.1). However, more often the statistically more robust relation (derived in [299])

$$w(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)} \tag{8.2}$$

is used, where DD is the number of observed galaxy pairs with an angular separation in the range $(\theta, \theta + \delta\theta)$, RR is the number of pairs for the same range of separations in a random distribution, and DR is the number of pairs with the same separation between the galaxies of the observed distribution and a random distribution.

In the local universe $w(\theta)$ is known to follow approximately a power law

$$w(\theta) \propto \theta^{-\gamma} \tag{8.3}$$

where (depending on the sample) γ ranges between about 0.7 and 0.8. This relation gives a reasonably good fit for all angles up to about 1 degree (e.g., [109]). For larger angles w decreases faster, and at much larger angles the local galaxy distribution cannot (within the statistical uncertainties) be distinguished from a uniform distribution. Precise derivations of the angular correlation function show also a small, but statistically significant, deviation from a power law at small angles (e.g., [577]).

The deeper cause of an angular two-point correlation function $w(\theta) \neq 0$ obviously is the non-uniform distribution of the galaxies in 3-D space. Therefore, the function $w(\theta)$ is related to the spatial two-point correlation function $\xi(r)$, which has been defined by Eq. (3.38) in Sect. 3.2.3. If the change of the galaxy distribution with distance and the cosmic parameters are known, $\xi(r)$ can be derived from $w(\theta)$ (e.g., [152, 205]). The relation between $w(\theta)$ and $\xi(r)$ was investigated first by D. N. Limber [315]. Therefore, conversions between these two quantities are called "Limber transformations." A detailed analysis shows that an angular two-point correlation function of the form

$$w(\theta) \propto \theta^{-\gamma} \tag{8.4}$$

transforms into a spatial two-point correlation function

$$\xi(r) = (r/r_0)^{-(\gamma+1)} \tag{8.5}$$

where r_0 is called "correlation radius," "correlation length," or "clustering scale." In the local universe the correlation length is about $r_0 = 5h^{-1}$ Mpc ≈ 7 Mpc. Correlation lengths (and two-point correlation functions) of the dark matter halo distribution can be predicted as a function of redshift from CDM simulations (e.g., [536]). Such data can be used to predict the two-point correlation function of galaxies, if the occupation of the dark-matter halos by the galaxies is understood and known.

Derivations of the angular two-point correlation function and of the corresponding spatial two-point correlation function have been carried out for many different

samples of high-redshift galaxies with redshifts in the range $2 < z < 4.5$ (e.g., [3, 279, 423, 573]). The observed correlation lengths and exponents γ are of the same order as those observed in the local universe. However, as first shown by M. Ouchi and colleagues [396], the two-point correlation function of Lyman-break galaxies at $z \approx 3$ cannot be fitted by a single power law. For angles $\theta > 7$ arcmin a power law with γ near 0.7 is found to provide a good fit to the data. However, for smaller angles the correlation function exceeds the power-law fit by large factors. In terms of the dark-matter models which assume that all galaxies reside in dark halos, the larger correlation found at small angles (corresponding to sizes of about $\leq$0.3 Mpc) can be explained by a transition between a single and a double occupancy of the halos, i.e., the occurrence of two galaxies in the same dark halo.

Another interesting result of the studies of the two-point correlation function at high redshift is a clear dependence of r_0 on the galaxy type and on the galaxy luminosity. According to R. Quadri and colleagues [423], galaxies meeting the "distant red galaxies" (DRG) criterion (see Sect. 7.2) show a correlation length which is about twice as large as that of other NIR selected galaxies at $2 < z < 3.5$. Moreover, as shown in Fig. 8.2, the observed r_0 depends clearly on the color selection criteria, with redder galaxies showing stronger clustering. Yoshida et al. [573] find the correlation length to depend strongly on the (rest-frame) UV and optical luminosity. Optically bright galaxies at $z \approx 3$ always have large correlation lengths. Since the optical luminosity (originating mainly from old stars) is a measure of the stellar mass of the galaxies, this indicates that massive galaxies always reside in massive halos. For galaxies which are faint in the optical bands, the correlation length is

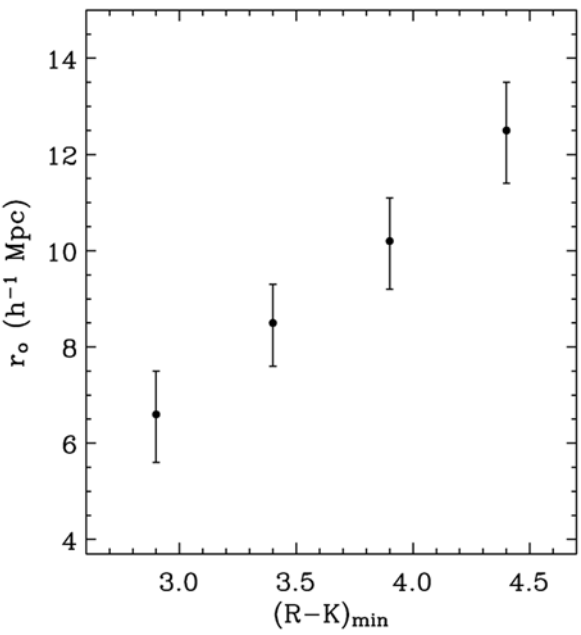

Fig. 8.2 The mean correlation length r_0 of samples of high-redshift galaxies, selected according to their minimum $R - K$ color index. The figure shows the steady increase of r_0 with increasingly red colors. (From Quadri et al. [423])

found to increase with the UV luminosity. Obviously in these systems the young stars dominate the UV flux and the mass.

The observed spatial two-point correlation functions have also been used to derive biasing factors (as defined in Eq. 3.37) for the corresponding samples. For this purpose b usually is calculated from the ratio

$$b = \sqrt{\frac{\xi_{galaxies}(r)}{\xi_{matter}(r)}} \tag{8.6}$$

where for ξ_{matter} the spatial two-point correlation function predicted by CDM models is assumed. Because of the dependence of r_0 on the galaxy type and on the luminosity, the resulting biasing factors also depend on these parameters. Some of the results are summarized in Fig. 8.3. The dashed lines in this figure indicate the theoretically predicted evolution with redshift of a given bias value. These predictions are based on the model of J. N. Fry [179], which takes into account the decrease of a given bias with time due to the increasing clustering of the dark matter. The model does not include effects of the evolution of the baryonic matter with time. Figure 8.3 is taken from Quadri et al. [423], where further details are explained.

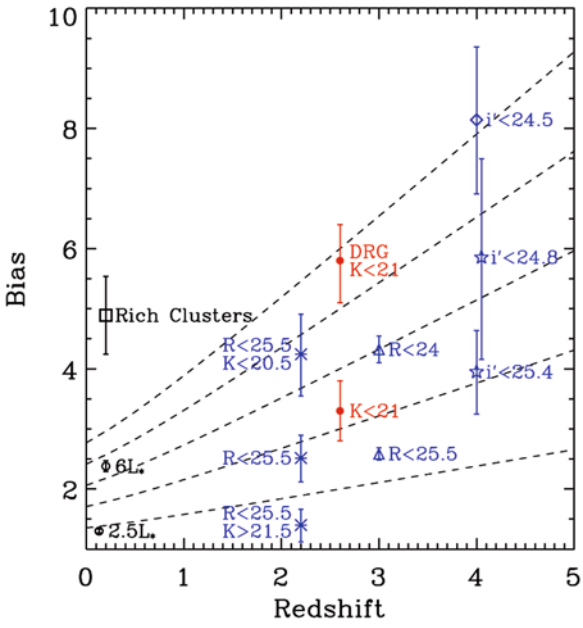

Fig. 8.3 Bias values of samples of high-redshift galaxies with different selection criteria. The *dashed lines* indicate the expected evolution of the bias with redshift. All bias values were determined from observed two-point correlation functions. (From Quadri et al. [423])

As shown by Fig. 8.3, the galaxies observed at high-redshift clearly show a higher biasing than the local galaxies. Moreover, at a given redshift the biasing increases with the luminosity.

The strong biasing of the high-redshift galaxies also explains why these galaxies do not follow the redshift dependence of r_0 predicted for dark halos. As shown in

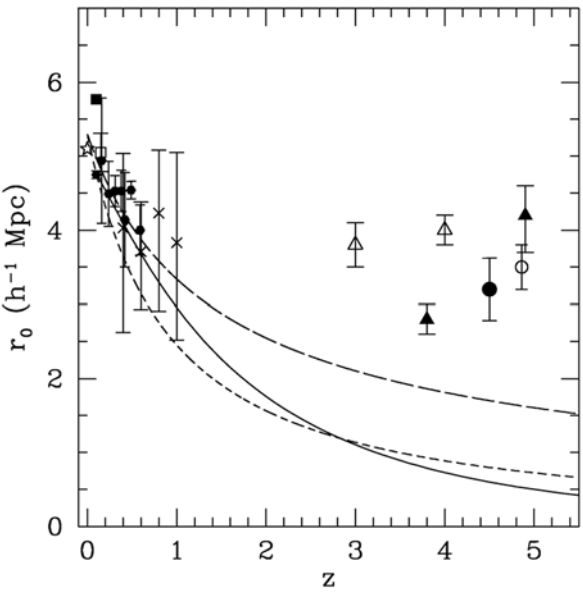

Fig. 8.4 The correlation length observed for local galaxies and for different samples of UV-bright high-redshift galaxies. The *triangles* correspond to samples of Lyman-break galaxies. The *open* and *filled circles* corresponds to LAE samples. The *solid line* indicates the redshift dependence of the correlation length predicted for the dark halos of CDM models (from [365]). The *broken lines* indicate predictions based on the "ϵ model" of Groth and Peebles [213]. The observed r_0 values of the high-redshift galaxies are clearly above the theoretical predictions. The differences can be explained by the higher biasing factor of these objects. (From Kovač et al. [279], where references to the individual data points can be found)

Fig. 8.4, even for the UV-selected high-redshift galaxies the correlation length is above the halo prediction by large factors. Only the local galaxies with their smaller biasing factors are close to the halo prediction.

8.3 Progenitors of Galaxy Clusters at High Redshift

Galaxy clusters are gravitationally bound systems with typical dynamical masses of the order $10^{13} < M/M_{\odot} < 3 \times 10^{15}$ (Sect. 2.3). According to Fig. 3.12 such massive bound matter concentrations form at late cosmic epochs and are expected to be extremely rare at redshifts $z > 2$. Thus, galaxy clusters with properties as observed in the local universe are not expected to exist at high redshift. This is consistent with the fact that, so far, no galaxy cluster with clear evidence for a virialized density structure could be discovered at $z > 2$. On the other hand, regions of high overdensities of galaxies are observed in deep galaxy samples at redshifts up to $z = 5.8$ (e.g., [354, 396, 397, 578]). Some of these overdense regions (which often are referred to as "protoclusters") are associated with luminous QSOs or radio

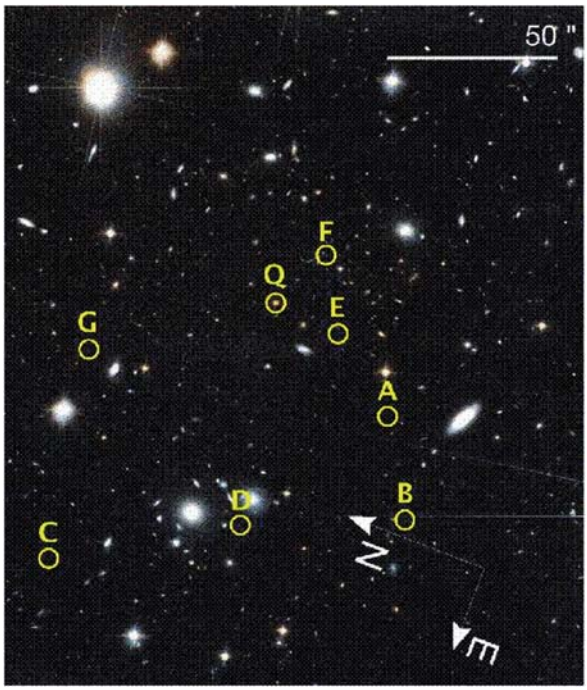

Fig. 8.5 HST ACS image of the field of the $z = 5.82$ QSO SDSS J0836+0054. The *circles* indicate the positions of the QSO (Q) and of 7 galaxies (A to G) which according to their colors have redshifts > 5.5. From the known surface density of galaxies matching the corresponding color selection only between 1 and 2 such galaxies were expected to be present in this image. (From Zheng et al. [578])

galaxies. An example of an apparent protocluster associated with a QSO is reproduced in Fig. 8.5.

In some cases the observed surface number densities of galaxies have been used to estimate the mass of the overdense regions and the spatial (3-D) overdensities. These studies resulted in total-mass values of the order $10^{14} M_{\odot}$. The reported galaxy number contrasts $\delta n/\bar{n}$ range between 2 and 80. The largest values are almost as large as the number density contrasts of present-day galaxy clusters. It seems almost certain that at least part of the observed overdense high-redshift regions are developing toward bound galaxy clusters of the present-day universe.

An important difference between the high-redshift protoclusters and the present-day galaxy clusters is their stellar population content. As explained in Sect. 2.3, the massive galaxy clusters observed in the local universe consist mainly of passively evolving galaxies, including massive ellipticals. The presence of luminous radio galaxies and QSOs in some of the protoclusters show that these systems include very massive galaxies. However, in all cases the overdensities and the total mass have been deduced from the distribution of UV-luminous star-forming galaxies or from LAEs. Thus, at least a major fraction of the protocluster members are starburst galaxies. The stellar masses of the individual starburst galaxies and LAEs are similar to those observed in the field for these types at the corresponding redshifts (e.g., [397]). Since these mass values are much smaller than the stellar mass of typical present-day cluster galaxies, frequent mergers are needed to built up the present day cluster ellipticals from the star-burst galaxies of the high-redshift

protoclusters. Whether the $z \approx 5$ protoclusters contain (in addition to the star-burst systems) already passively evolving galaxies cannot be determined unambiguously at present, since at the redshifts in question passively evolving galaxies cannot (yet) be identified reliably (e.g., [97]). However, while observing protoclusters around four radio galaxies with redshifts $2 < z < 3$ Kodama and colleagues [272] found that in addition to the known overdensity of blue (starburst) galaxies these protoclusters also show an overdensity of NIR-selected galaxies with overdensity factors of 2 to 3 relative to the general field. From these observations they conclude that in these protoclusters there exists already the bright end of a red sequence of passively evolving galaxies. This nascent red sequence was found to be more pronounced at $z \approx 2$ than at $z \approx 3$. Future observations with more efficient and more sensitive infrared instruments will probably allows us to derive more detailed information on the population content of the protoclusters.

Chapter 9
Interactions with the Intergalactic Medium

9.1 Properties of the High-Redshift IGM

In Chap. 2 it has been pointed that most of the baryonic matter of the universe is found in the intergalactic space (see, e.g., Table 2.1). Galaxies are known to interact with this intergalactic matter, and these interactions have important consequences for the IGM as well as for the galaxies. The interactions and their effects depend on the properties of the intergalactic matter. But in spite of its predominance of the baryonic mass, the intergalactic gas remains one of the less well known components of the cosmic matter. In the following paragraphs our limited knowledge on the high-redshift IGM properties will be briefly summarized.

9.1.1 The Intergalactic Gas After Recombination

The evolution of the baryonic matter after the recombination of the initial cosmic plasma, and after the decoupling of the gas and the radiation field, has been described in Sect. 3.3. When the recombination was complete, the cosmic gas continued to cool with the cosmic expansion. After recombination the absorption cross section of the neutral gas for CMB photons was very small. However, since there were $>10^9$ more photons than atoms, the absorption of CMB photons still had a significant effect on the gas (while the effect of the gas on the radiation field was negligible). The occasional absorption of CMB photons kept the kinetic temperature of the expanding gas in equilibrium with the CMB until a redshift of $z \approx 200$ was reached. Thus, up to this cosmic age the gas cooled (like the CMB) $\propto (1+z)$. After that epoch heating by the CMB became negligible, the gas started cooling adiabatically, i.e., $\propto (1+z)^2$. With the formation of bound dark halos, the gas started accumulating in these halos. Due to the compression in the gravitational potential of the halos, the gas started to heat up again. However, because of the low mass of the first halos, initially the temperature of the gas remained relatively low, and practically all the gas remained neutral until the first stars and black holes formed. Only then the UV radiation of the stars and black-hole accretion disks started to reionize the IGM gas.

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_9,

During the adiabatic cooling phase the neutral hydrogen gas can, in principle, be observed by its H I 21-cm line absorption. Since at this phase the gas cools faster than the CMB (and since the H I spin temperature can be shown to follow the kinetic temperature), the gas is expected to produce locally a 21-cm absorption line in the CMB spectrum. If the hydrogen is homogeneously distributed, the continuous increase of the redshift along the line-of-sight converts this local line absorption into an observed smooth continuum absorption. If the medium is not homogeneous, the observed spectrum will show structure. Therefore, high redshift H I observations, in principle, provide an important tool to study the interstellar gas at these epochs.

Unfortunately, observations of the redshifted 21-cm line are technically rather difficult and demanding. For the redshifts of interest ($10 < z < 200$) the 21-cm line is shifted to the low-frequency radio range, which is affected by the terrestrial ionosphere and by strong backgrounds. Therefore, studies of the high-redshift 21-cm absorption (and emission) are outside the sensitivity of the present-day radio telescopes. But new instruments under construction or in an advanced planning phase will probably make such observations possible in the near future (see Sect. 12.1.1).

9.1.2 The Lyα Forest Absorbers

Most of our present knowledge on the high-redshift intergalactic medium is based on observations of the Lyman-series absorption lines of hydrogen and the associated metallic lines in the spectra of high-redshift QSOs. Most of these Lyman-line absorption systems have neutral hydrogen column densities $N(\mathrm{H\ I}) < 10^{17}\,\mathrm{cm}^{-2}$. These systems are referred to as the "Lyα forest" or "Lyman forest." Absorption systems in QSO spectra with higher column densities are discussed in Sect. 9.4.

The basic properties of the Lyα forest lines have already been discussed in Sect. 6.9. (For a comprehensive review of this subject see [424].) The individual Lyα forest lines correspond to regions of an enhanced density of hydrogen atoms along the line of sight to the background source. At the epochs at which these density enhancements were formed, the gas pressure effects were still small. Thus the Lyα absorbing gas is expected to show essentially the same distribution as the dark matter. This is confirmed by hydrodynamic simulations of the evolution of the gas (see Fig. 9.1). The density of the gas producing the forest lines is typically about 10 times higher than the mean IGM gas density at the corresponding epoch. The temperature of the Lyα absorbing gas is of the order 10^4 K. The mean metallicity of the Lyman forest gas varies between about 0.001 $Z_\odot$ and 0.0001 $Z_\odot$ [485].

Since (at present) QSOs are known up to $z = 6.4$, the Lyman forest lines allow us to probe the neutral hydrogen distribution up to that redshift. At the epochs corresponding to this redshift, almost all the hydrogen has already been ionized by the UV radiation of the galaxies and QSOs. Thus, the Lyman absorption lines are typically produced by ionized gas concentrations with a small fraction of neutral hydrogen atoms.

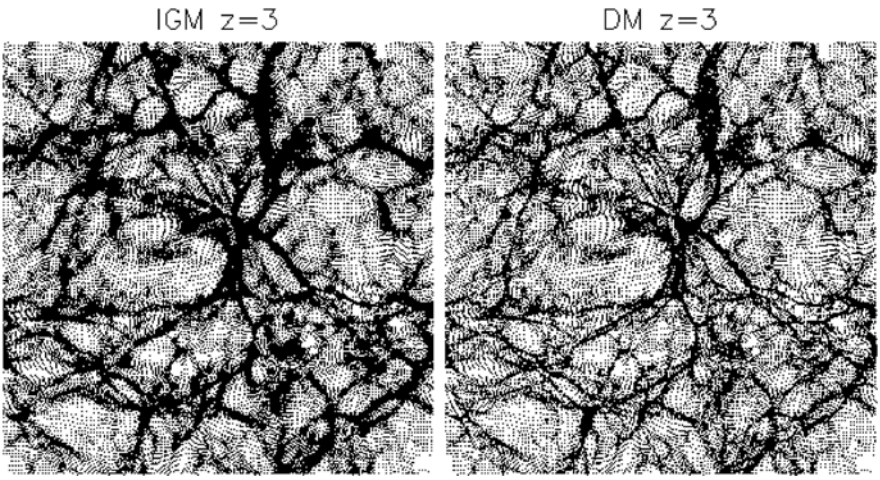

Fig. 9.1 Results of hydrodynamic simulations of the IGM gas distribution and of the dark-matter distribution at redshift $z = 3$. The high-density regions are indicated in *black*. The (comoving) size of the space slices reproduced in the figure is 12 Mp/h on a side and 0.1 Mpc/h in depth. Apart from somewhat larger widths, the gas filaments follow closely the DM distribution. (From Viel et al. [547])

In Sect. 6.9 it has been pointed out already that the properties of the Lyman forest changes rapidly with cosmic time. The evolution of the IGM transmission of the Lyman forest between $z = 2$ and $z = 6.3$ has been investigated in detail by A. Songaila [486], who finds a smooth change of the transmission properties of the gas over the whole redshift range, indicating a smooth variation of the ionization rate during the corresponding cosmic time interval.

As noted above, the Lyman forest lines are produced in a mostly ionized gas. In fact, the Lyα forest of distant QSOs could not be observed if the IGM were neutral. Due to the high absorption cross-section of H I for Lyα light, all radiation emitted by the distant QSOs shortward of the QSO's Lyα line would be completely absorbed by a neutral IGM. The complete absorption of the flux shortward of the Lyα line of a high-redshift radiation source in a neutral IGM was first predicted by J. E. Gunn and B. A. Peterson in 1965 [215] and is known as the "Gunn–Peterson effect." The broad absorption features with zero flux caused by the Gunn–Peterson effect are called "Gunn–Peterson troughs." Their presence obviously provides a tool for determining the epoch and redshift when neutral IGM started to become reionized by the first galaxies and QSOs.

Figure 9.2 shows that in the spectra of QSOs with $z > 5.8$ indeed very little flux is observed shortward of the QSO's Lyα line. In the spectrum of the QSO with the largest redshift included in Fig. 9.2 (J103027.10+052455), no significant flux is seen just shortward of the Lyα emission line. This has been interpreted as a Gunn–Peterson trough in this object. However, a practically zero flux can also be produced by overlapping Lyα lines, which could not be separated at the resolution

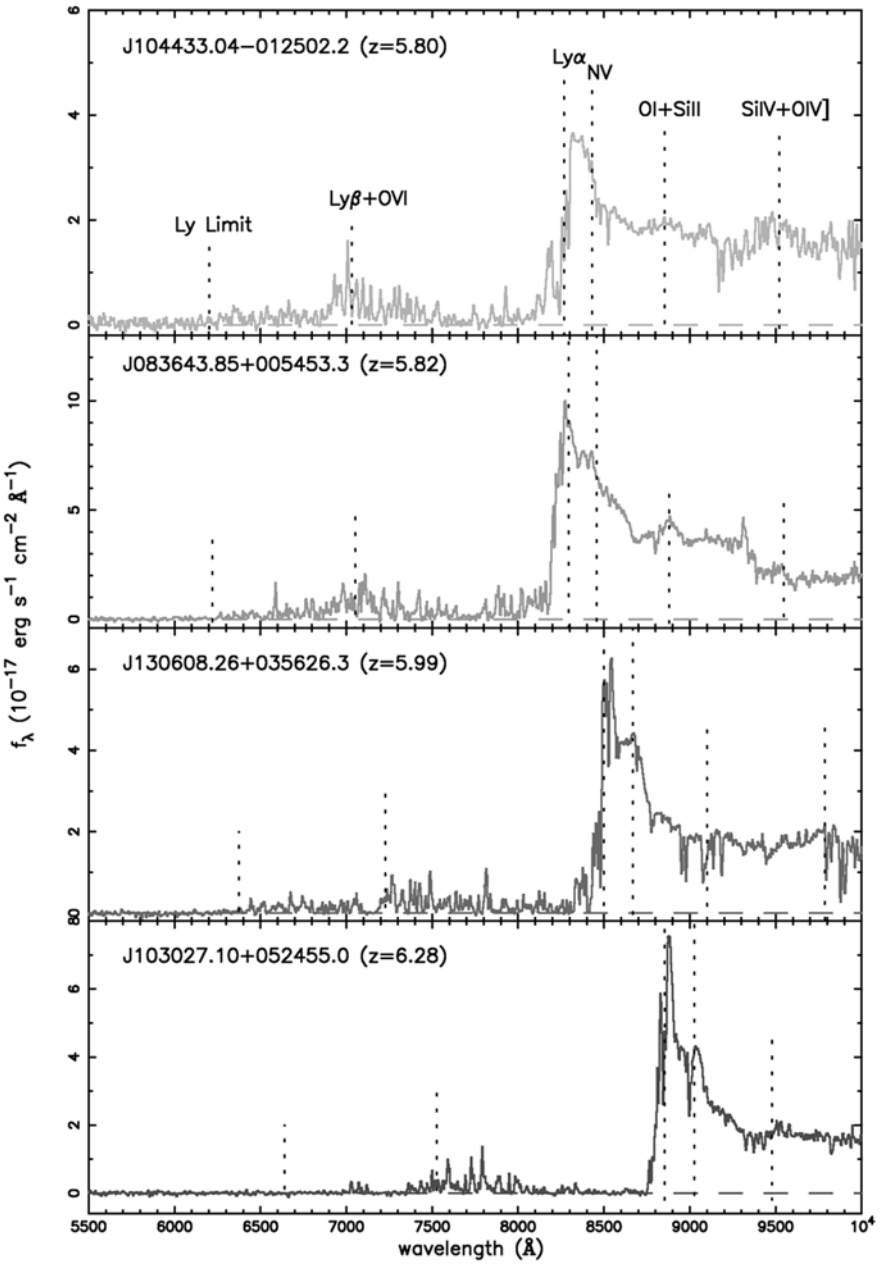

Fig. 9.2 Keck spectra of four high-redshift QSOs. Note the increase with redshift of the absorption of the flux shortward of the Lyα line of the QSOs. (From Becker et al. [33])

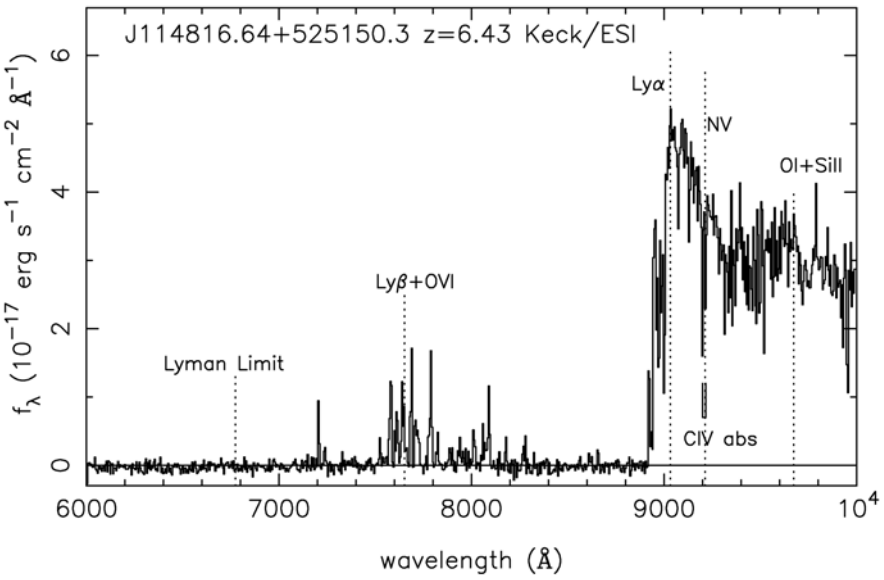

Fig. 9.3 Spectrum of the distant QSO J1148+5251 ($z = 6.42$). Note the presence of spikes of light transmission shortward of the Lyα line of the QSO. (From Fan et al. [167])

of the spectra of Fig. 9.2. Moreover, in the spectrum of the higher redshift QSO, which is reproduced in Fig. 9.3, some flux is observed shortward of the Lyα emission line. A careful evaluation of the spectra of the QSOs known in the redshift range $5.7 \leq z \leq 6.4$ by Gallerani and colleagues [194] shows that at the cosmic time corresponding to the redshift range $5.3 < z < 5.6$ the IGM was already highly ionized ($N(HI)/N(H) \approx 10^{-4.3}$). But for $z = 6.3$ only an upper limit $N(HI)/N(H) < 0.36$ could be derived. According to Gallerani et al., the wavelength regions of high transmission in the spectra of the very-high redshift QSOs can be identified with overdense regions and H II regions around starburst galaxies and QSOs. A narrow transmission peak in the spectrum of the QSO SDSS J1148+5251 ($z = 6.42$, see Fig. 9.3) could be identified as a IGM region fully ionized by the intervening QSO RD J1148-5253 at $z = 5.70$. The size of the H II region of the foreground QSO was found to be about 39 Mpc, and the minimum lifetime of this QSO could be estimated to 11 Myr [194].

9.1.3 Hot IGM Plasma at High Redshift

In the local universe much of the intergalactic matter is observed as a hot ($\approx 10^7$ K) plasma residing in the deep potential wells of massive galaxy clusters. The high temperature is due to the gas being in hydrostatic equilibrium in the deep potentials of the clusters. Since most of the dynamical mass of the galaxy clusters is dark

matter, the hot gas has actually been captured by the massive dark halos of the clusters.

As illustrated by Fig. 3.12, the present-day massive dark halos of galaxy clusters formed at rather late cosmic epochs, and massive halos were rare at the epochs where the high-redshift galaxies are observed. As noted above, the temperatures resulting from the gas accumulation in the early low-mass halos were much lower. Therefore, at the epochs of the high-redshift galaxies, hot hydrostatic inter-cluster gas can be neglected.

A hot intergalactic plasma component can also form as a result of the X-ray and FUV radiation of the first galaxies and QSOs. This hard radiation is expected to heat up the low-density gas, which cannot cool efficiently by radiation. If a hot plasma produced by this mechanism exists, it is expected to produce bremsstrahlung at X-ray wavelength. Observationally, this radiation would occur as a diffuse X-ray background. Moreover, the free electrons of a hot intergalactic plasma at high redshift are expected to produce a strong signature in the CMB polarization.

None of these predictions is supported by the observations. The cosmic X-ray radiation field can be fully accounted for by known discrete sources [70], and the observed large-scale CMB polarization can be explained by the warm ionized gas produced during reionization. Thus, at present, there is no observational evidence for the existence of a significant amount of a hot IGM plasma at redshifts $z > 2$.

9.2 Impact of the High-Redshift Galaxies on the IGM

According to the theoretical concepts summarized in Sect. 3.3, the formation of galaxies started in a neutral gas which consisted of hydrogen and helium. In the present-day universe the IGM is almost fully ionized ($N(HI)/N(H) \approx 10^{-5}$) and the IGM gas has been significantly enriched with heavy elements. Obviously, the formation of stars and galaxies has profoundly influenced the properties of the IGM. The main processes which caused these changes of the IGM properties are summarized in this section.

9.2.1 Radiative Effects

9.2.1.1 The Basic Mechanisms

The most obvious effect of the galaxies on their environment is the energy input by radiation. But not all the radiation emanating from galaxies affects the IGM. In terms of total energy, the light emitted by the stars dominates the energy output of galaxies. However, most of the radiation of the evolved stars is emitted at visual and IR wavelengths, to which the IGM is essentially transparent. Only the UV flux of the hot massive stars is strongly absorbed by the neutral hydrogen atoms of the IGM. Also absorbed are the energetic photons produced in galaxies by the AGN, by stellar X-ray sources, and by supernova remnants.

Apart from heating the IGM, the absorption of X-ray photons and of UV photons with $\lambda < 912$ Å results in the ionization of the neutral fraction of the intergalactic gas. Under normal conditions the ionizing effect of the UV photons is more important than that of the X-rays. On the other hand, the hard X-ray photons reach deeper into the space and affect larger volumes.

The effect of the UV radiation from starburst galaxies on the IGM obviously depend on the space density or luminosity function of these objects, and on the fraction of the UV radiation which escapes the galaxies. Our present knowledge on the luminosity function of the high-redshift starburst galaxies has been summarized in Sect. 7.1. In that section it has also been explained that much of the UV radiation produced by the hot stars of star-forming galaxies is absorbed inside these systems. Therefore, the effect of the high-redshift starburst galaxies depends critically on the fraction of the $\lambda < 912$ Å flux which actually escapes from the galaxies. In principle, this fraction can be measured from spectra which include the rest-frame EUV spectral flux. After correction for the Lyman-continuum absorption by the IGM and the galactic interstellar matter, the emitted EUV flux can be determined from the spectra, and the fraction of escaping EUV radiation can be derived by comparing the observed spectra with model spectra.

Figure 9.4 shows a high-S/N rest-frame FUV-EUV composite spectrum of $z \approx 3$ starburst galaxies. As shown by this figure, as a result of the intergalactic absorption, the observed flux shortward of 912 Å is very weak. From the spectrum of Fig. 9.4, Shapley et al. [472] found an observed flux at 900 Å which is about 1.7% of the flux

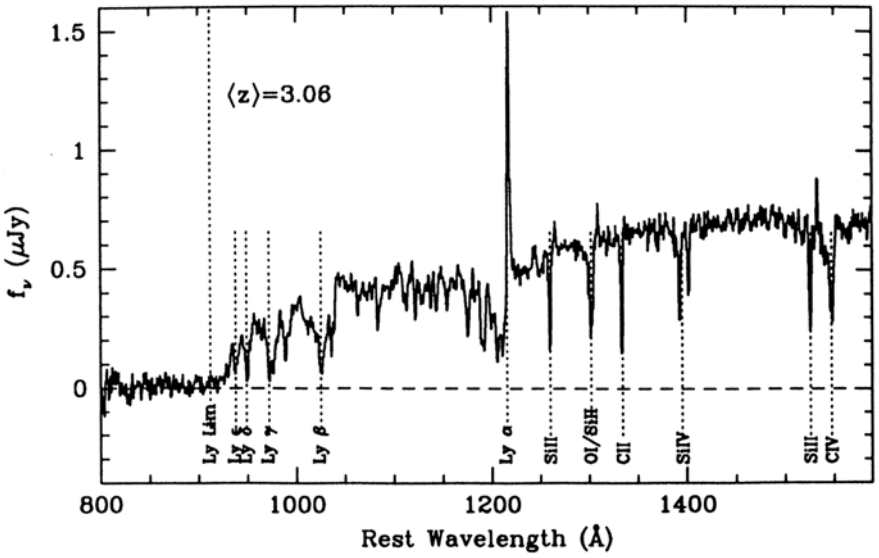

Fig. 9.4 Rest-frame FUV-EUV composite spectrum of $z \approx 3$ Lyman break galaxies. A small, but statistically significant, Lyman continuum flux is present in this spectrum. (From Shapley et al. [472])

at 1500 Å. A correction for IGM absorption (assuming an average IGM opacity) rises this percentage to about 4.5%. Assuming a plausible intrinsic ratio of ionizing and non-ionizing UV flux, this number converts into an escape fraction of the ionizing radiation of about 14%. However, as pointed out by Shapley and colleagues, it is not clear whether these data are representative for the starburst galaxies at this redshift. Moreover, within the sample of 14 starburst galaxies investigated by Shapley et al. there are several objects with much higher than average Lyman continuum emission. The spectra of two of these objects are shown in Fig. 9.5. For these two galaxies the corrected flux at 900 Å is between about 22% and 34% of the flux at 1500 Å, i.e. 5–8 times larger than the average. Because of these large variations of the amount of escaping UV radiation, the results derived from small samples have to be taken with caution.

A photometric search for the Lyman continuum emission of Lyman-break galaxies and LAEs with redshifts $z \approx 3.1$ has been carried out by Iwata and colleagues [247], using the Suprime camera at the Subaru telescope. From their data these authors estimate an average escape fraction of the Lyman continuum photons $>15\%$.

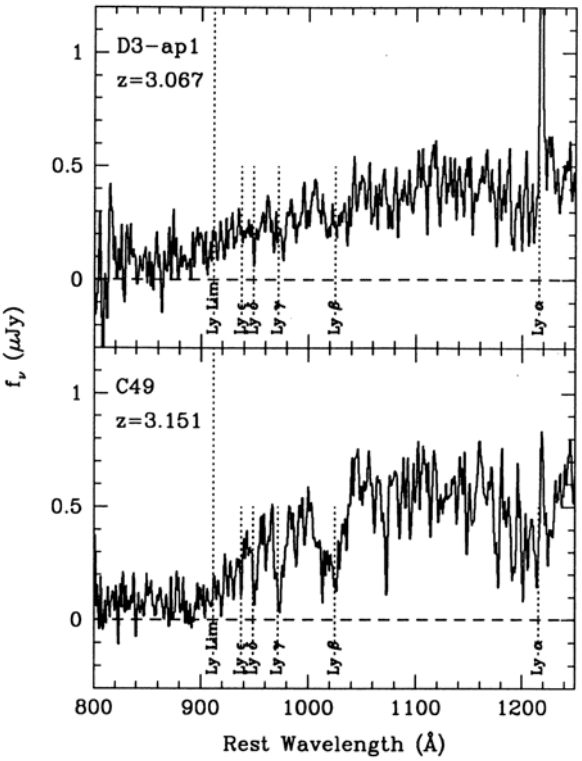

Fig. 9.5 Examples of individual high-redshift starburst galaxies with significant flux in the Lyman continuum. (From Shapley et al. [472])

There are no good data on the Lyman continuum emission of higher redshift objects. As pointed out in Sect. 7.1, with increasing redshift internal dust extinction becomes less important, the galaxies tend to be smaller, and higher redshift galaxies tend to have lower metallicities. As a result their young stars are expected have higher average surface temperatures. Therefore, it appears plausible to expect that with increasing redshift the ratio between the flux at 900 Å and at 1500 Å becomes larger. Future data will hopefully allow us to verify this assumption.

The luminosity functions of the AGN and QSOs are well established up to redshifts of about $z = 3$. Their spectra are known to vary little with the redshift. Therefore, their contribution to the UV radiation field at low and intermediate redshifts can be estimated reliably (e.g., [245]). Because of uncertainties concerning the luminosity function at high redshifts, the contribution of the high-redshift QSOs is less clear. Some information on the QSO LF at $z \approx 6$ could be gained recently from the identification of intrinsically relatively faint $z > 5$ QSOs in the "Sloan Digital Sky Survey Deep Stripe" (SDSS II) survey [252]. These data suggest a rather small contribution of the AGN to the high-redshift EUV radiation field.

9.2.1.2 The Cosmic Reionization

Closely connected with effects of the UV radiation of the high-redshift galaxies on the IGM is the reionization history of the universe. As pointed out in Sect. 9.1.2, studies of the Lyα forest and the (missing) Gunn–Peterson effect indicate that at a redshift of $z = 5.6$ the IGM was already highly ionized and that at $z = 6.3$ at most about 1/3 of the hydrogen was still neutral [194]. These data indicate that the reionization of the IGM occurred before the epoch corresponding to $z = 6$.

Another constraint on the epoch of reionization is provided by the large scale polarization of the CMB. This polarization is produced by Thomson scattering of the CMB by the free electrons of the inhomogeneous reionized medium. If the degree of ionization as a function of redshift is known, the polarization can be predicted from the observed temperature variation of the CMB (e.g., [574]). Vice versa, the ionization as a function of redshift can be inferred from the observed polarization. Since the polarization signal is damped on scales smaller than the horizon at the epoch of reionization, the reionization affects mainly the CMB polarization at large angular scales. A detailed discussion this subject (and other aspects of the cosmic reionization) can be found in a review presented by Fan et al. [165]. Observational results on the CMB polarization have been presented, e.g., by Spergel and colleagues [491]. According to these authors, the CMB polarization derived from the WMAP 3-year data indicates a start of the reionization at an epoch corresponding to $z \approx 11$.

Other tracers of the epoch of reionization are the observations of LAEs at high redshift, and the size of the ionized IGM volumes (H II regions) around high-redshift QSOs. In a fully neutral IGM the Lyα emission lines of LAEs are expected to be completely absorbed by the red damping wings of the Gunn–Peterson troughs. Thus,

observations of LAEs up to $z = 6.96$ [393] shows that a significant amount of reionized gas must have been present already at the epoch corresponding to $z \approx 7$.

The size-dependence of the H II regions around QSOs (and starburst galaxies) results from the fact that in a given time interval the central radiation source must provide the photons to ionize all hydrogen atoms inside the corresponding volume. In first approximation the radius r_{HII} of the H II region around a QSO (or galaxy) is given by

$$r_{HII} = \left(\frac{3\dot{N}_p t}{4\pi n_{HI}} \right)^{1/3} \tag{9.1}$$

where $\dot{N}_p$ is the rate of ionizing photons of the radiation source, t is the time since the onset of the radiation, and n_{HI} is the density of the neutral hydrogen. Thus, for an observed QSO luminosity and plausible assumptions on the QSO lifetime, the average size of the H II regions is directly related to the neutral hydrogen density. Observationally, the H II regions around QSOs normally are detected by the absence of H I absorption at redshifts slightly smaller than the QSO emission line redshift. In the astronomical literature this lack of H I absorption close to the emission-line redshift is referred to as the "proximity effect."

Using the H II region sizes inferred from Lyman forest observations, lower limits of the neutral hydrogen fractions between 0.03 and 0.1 at $z \approx 6$ have been estimated. However, only few suitable QSOs are known at this high redshift. Consequently, these values are based on very small samples. As pointed out by Bolton and Haehnelt [55], because of the density fluctuations of the IGM gas, much larger samples of high-redshift H II regions of QSOs are needed to constrain the neutral-hydrogen fraction reliably by this method.

Simple estimates show that the present-day UV radiation field of the galaxies and QSOs can easily maintain the high ionization fraction of the IGM. However, the UV radiation field at high redshifts is not known well enough to determine the epoch and the details of the reionization from the formation history of galaxies and AGN. Nevertheless, it seems unlikely that the QSOs were responsible for the cosmic reionization, since their space density at redshifts ≥ 6 appears to be too low to provide the necessary UV or X-ray flux (e.g., [252]). On the other hand, within the limits of our present knowledge on their luminosity functions and spectral properties, the high-redshift starburst galaxies could probably have produced the UV photons required to cause and to maintain the ionization fraction observed at $z = 6$ (e.g., [56, 65]).

As pointed out in Chap. 3 the formation of stars and galaxies begun in volumes of exceptionally high cosmic density, and then spread to the lower density areas. Therefore, we have to assume that the reionization of the IGM occurred over an extended period of time, starting in high-density regions, where galaxy formation begun early. As soon as the first massive young stars started producing ionizing radiation, H II regions formed around their host galaxies. With time the size and the number of these ionized volumes became larger. Finally, the individual H II regions

around the galaxies started to merge, leaving only filaments of neutral matter, which then were ionized by the general IGM radiation field. As a result of the energy input into the IGM gas, the gas concentrations in low-mass dark halos may have become heated above their virial temperatures. In this way the gas concentrations in the least massive halos may have been evaporated before star formation could set in [31]. This may explain why the space density of galaxies of very low mass is significantly lower than the predicted density of low-mass dark halos.

9.2.2 Galactic Winds

Among the characteristic properties of high-redshift starburst galaxies discussed in Sect. 7.1 are massive outflows from these objects. In the observed spectra the outflows produce blueshifted absorption lines which indicate outflow velocities up to several $10^2\,\mathrm{km\,s^{-1}}$ (e.g. Fig. 7.30). In view of the modest masses of the UV-bright high-redshift starburst galaxies, much of the outflowing matter is expected to be lost from these systems. The energy and momentum sources of these outflows are mainly supernovae and stellar winds. For the $z = 2.727$ starburst galaxy MS 1512-cB58, Max Pettini and colleagues estimated a mass loss rate of 60 $M_\odot\ \mathrm{yr}^{-1}$. This is of the same order as the star formation rate of this galaxy. If the starburst phase continues for 10^8 yr, about $10^{10} M_\odot$ of interstellar matter are transported into the IGM from this galaxy. Shock waves produced by the outflows can contribute to the ionization of the environment of the young starburst galaxies [4]. But, as shown by hydrodynamical model computations, the most important effect of the galactic winds is the chemical enrichment of the IGM gas (see the next section). The outflows may also suppress the formation of low-mass galaxies close to massive systems. On the other hand, according to hydrodynamic models, the Lyman forest clouds and the statistics of the Lyman forest H I lines are little affected by the outflows [517].

Practically, all our information on the winds from high-redshift galaxies is based on objects with $z \approx 3$. So far no direct empirical information is available for higher redshift galaxies. But from physical arguments it appears likely that mass loss plays an even more important role at higher redshifts. According to the hierarchical model of galaxy formation, at the epochs corresponding to very high redshifts the galaxies and their dark-matter halos were less massive and their potential wells less deep than at $z = 3$. On the other hand, the specific energy and momentum produced by SNe and stellar winds in individual young starbursts was probably as high as at lower redshifts. Thus, at redshifts $z > 3$, the winds were probably even more efficient in transporting matter into the IGM. Therefore, much of the chemical enrichment of the IGM may have occurred at a very early epochs (e.g., [326]). As noted below, this is supported by the observed chemical abundances of the IGM as a function of redshift.

9.2.3 The Chemical Enrichment of the IGM

As noted in the preceding section, one of the most important effects of the mass outflows of the high-redshift galaxies is the enrichment of the IGM gas with heavy elements. The resulting enrichment history of the IGM has been investigated using observations of the lines of elements heavier than helium in Lyman-forest systems. These observations have been supplemented by hydrodynamic model computations of the enrichment process. The observations have shown that elements heavier than helium are present in a majority (>75%) of the observed Lyman-forest clouds [488]. However, the enrichment differs between individual clouds, and there are Lyman-forest systems where no trace of heavy-element absorption could be detected so far.

The well-investigated Lyman-forest systems contain resonance absorption lines of many different elements. The observed line ratios are in reasonable agreement with the theoretical predictions based on galactic winds (e.g., [426]). For some elements it is possible to estimate the mass of the absorbing material. Among the most prominent absorption feature in Lyman-forest systems is the C IV 1550 Å resonance doublet. Therefore, this doublet is often used as a measure of the chemical abundance of the Lyman-forest clouds. The C IV feature is always strong, since under the physical conditions in the Lyman-forest clouds much of the carbon is present as C^{3+}. As noted in Sect. 9.1.2, the observed carbon abundance of the Lyman-forest clouds (derived from the C IV resonance doublet) corresponds to metallicities between about 0.001 $Z_{\odot}$ and 0.0001 $Z_{\odot}$. In Fig. 9.6, the observed CIV abundance is plotted as a function of the redshift of the corresponding clouds. Unexpectedly in the redshift range $2 < z < 6$ no statistically significant evolution of the C IV abundance with redshift can be detected (e.g., [477, 485, 487]). Obviously much of the chemical enrichment of the Lyman-forest clouds has already taken place during the first Gyr of cosmic time. According to A. Songaila, the fraction of the cosmic mass, which (at least since $z \approx 5$) is present in the form of C^{3+} in the IGM, is $\Omega_{CIV} \approx 2 \times 10^{-8}$ [485].

Miralda-Escudé and Rees [356] have estimated that the carbon production of the stars which were required to reionize the IGM results in a total cosmic metallicity of $Z = 10^{-5} Z_{\odot}$. Since most of the cosmic baryonic matter resides in the IGM (Sect. 2.7), this number can be directly compared with the observed CIV abundance in Lyman-forest clouds. The comparison seems to confirm that the stellar radiation of $z > 6$ starburst galaxies provided all the photons needed to reionize the IGM.

9.3 Accretion of IGM Gas

In the preceding sections the effects of the high-redshift galaxies on the IGM have been described. However, the IGM also affects the evolution of the galaxies. Among the important processes influencing the galaxies is the infall of matter from the IGM. This matter can originate in the same halo, or it can be gas of a small neighboring halo which is accreted by the galaxy's halo. As pointed out above, the IGM is not

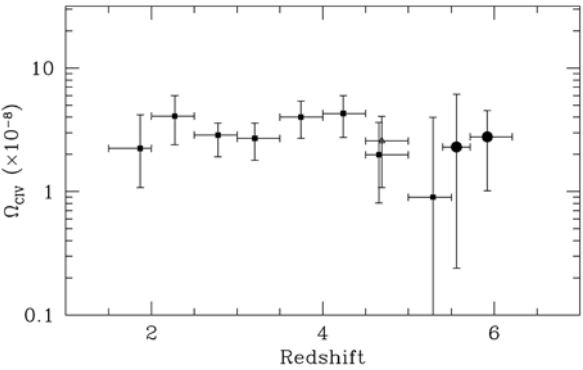

Fig. 9.6 Observed CIV abundance in Lyman-forest clouds as a function of redshift. The CIV abundance is expressed by the corresponding fraction of the total cosmic mass. The *filled squares* indicate data points from [485]. The *filled circles* are from [477], and the *triangle* from [414]. (From Simcoe [477])

free of heavy elements. However, its metallicity is much smaller than that of the ISM of a galaxy. Therefore, regardless of its exact origin, accreted matter always has a lower metallicity. Hence, the accretion of IGM matter affects the chemical evolution of galaxies.

Hydrodynamic model computations show that there are two ways that galaxies can accrete gas (e.g., [266]). If the infalling gas is decelerated in shock fronts, it is typically heated to temperatures of the order 10^6 K (for a Milky Way like galaxy). This hot gas then cools slowly to reach the temperature of the interstellar medium of the galaxies. Alternatively the gas can radiate the energy gained by compression at lower temperatures ($\approx 10^5$ K). This second accretion mode is usually referred to as the "cold mode." This mode develops if the gas cooling time is shorter than the local dynamic time. According to the model computations the cold mode dominates the mass accretion of the low-mass galaxies (with a baryonic mass less than about $\approx 2^{10} M_\odot$) and of the high-redshift systems. The hot mode dominates the mass accretion in groups and clusters at low redshift.

Since the density of the IGM gas is decreasing with cosmic time, the mass accretion of the galaxies becomes smaller at low redshifts. The decline of the cold-mode accretion directly follows the decreasing density. The accretion via the hot mode is further reduced by the increase of the cooling time with decreasing density.

As a result of the hydrodynamic simulations, the basic physics of the mass accretion by the galaxies appears to be reasonably well understood. On the other hand, so far, there is little direct observational evidence from the high-redshift universe which would allow us to verify the theory.

9.4 Damped Lyα Systems

In Sect. 9.1.2 it has been pointed out that the spectra of distant QSOs occasionally show, in addition to the Lyman-forest absorption lines (which are characterized by a neutral-hydrogen column density $N(\mathrm{H\ I}) < 10^{17}\,\mathrm{cm}^{-2}$), absorption line systems with higher column densities. These high-density systems are classified as "Lyman limit systems" (LLSs) if $10^{17}\,\mathrm{cm}^{-2} < N(\mathrm{H\ I}) < 10^{20}\,\mathrm{cm}^{-2}$ and as "damped Lyα systems" (DLAs) if $N(\mathrm{H\ I}) > 10^{20}\,\mathrm{cm}^{-2}$. While Lyman-forest systems are optically thin at the Lyman limit, the LLSs and DLAs are optically thick in the Lyman continuum. The damped Lyα systems differ from the Lyman-limit systems and (and the Lyman-forest systems) by the fact that the DLA gas is essentially neutral, the hydrogen of the two other classes of absorbers is highly ionized. (Thus, physically the LLSs are more closely related to the Lyman-forest systems). Observationally, the DLAs are characterized by very broad Lyα absorption lines with conspicuous damping wings (Fig. 9.7). A comprehensive review of the basic properties of the DLAs can be found in [565].

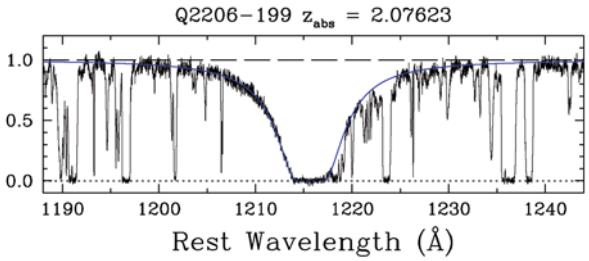

Fig. 9.7 Examples of a Lyα absorption line profile produced by a damped Lyα system (DLA). Note the much larger width of the DLA line as compared to the neighboring Lyman-forest absorption lines. (From Pettini et al. [412])

While the Lyman-forest lines and LLSs are formed in the IGM gas, the DLAs are assumed to result from absorption of QSO light by the interstellar medium of intervening galaxies [139]. In principle DLAs can be observed up to the distance where QSOs are known (i.e., $z \leq 6.4$). In practice, the redshift range is more limited since the crowding of Lyman-forest lines makes it difficult to evaluate DLAs at large redshift (see Fig. 9.2). Nevertheless, observations of DLAs provide an important tool to study the ISM of distant galaxies.

Among the important results derived from studies of the DLAs is their chemical composition. In general the relative abundances are similar to those in the Milky Way ISM (e.g., [412, 413, 565]). However, the DLA gas tends to show less evidence for dust depletion. Moreover, the mean metallicities of the DLAs are systematically lower than those observed in galaxies at the same redshift. The observed metallicities range between $10^{-3} Z_\odot$ and approximately solar values [565]. Like the chemical abundances observed in high-redshift galaxies (Sect. 7.1.2), the metallicity of the DLAs is decreasing with increasing redshift [565]. Detailed chemical abundance derivations for low-metallicity ($Z < 10^{-2} Z_\odot$) DLAs show C, N, and O abundances

similar to those observed in galactic (halo) low-metallicity stars [417]. The observed ratios between these elements indicate a formation by an early population of massive stars.

As noted above, DLAs in QSO spectra are assumed to be formed in the ISM of galaxies which happen to be located in front of the QSO. The high flux of the QSOs usually makes it difficult to detect these galaxies directly. Therefore, only a few high-redshift DLA host galaxies could be identified so far. (A list of the detected DLA hosts can be found in [565].) Among the known DLA host galaxies is a luminous Lyman-break galaxy (DLA 2206-19A, e.g., [362]). The other identified high-redshift DLA host galaxies are small and faint objects with properties similar to those of the Lyα-emitting galaxies (LAEs) described in Sect. 7.1.3. According to M. Rauch and colleagues the DLA host galaxies are identical with the population of faint extended LAEs detected in recent spectroscopic searches for low-surface brightness Lyα emitters [425].

Normally it is not known in which region of the galaxy the line of sight from the QSO is intersecting the ISM of the galaxy. This makes it difficult to derive reliable information on the structure of and physical conditions in DLA host galaxies. Future observations with high sensitivity and better angular resolution may help in the interpretation of these data.

Chapter 10
Implications

In the preceding sections the observed properties of the high-redshift galaxies have been described. There are still gaps in our knowledge, and some quantitative results have uncomfortably large error bars. Nevertheless, the observational results which have been accumulated during the past years already provide important information and valuable empirical constraints on the formation of galaxies and on the origin of the present-day structured universe. In the following some of these observational results will be compared to the theoretical predictions, and the implications of the observational facts for our present cosmological concepts will be discussed.

10.1 Comparison of Observations and Theoretical Predictions

Our present theoretical assumptions on the origin of cosmic structure and the formation of stars and galaxies have been outlined in Sects. 3.2 and 3.3. According to the theory described in these sections, galaxies are formed by the accumulation of baryonic matter in dark-matter halos at epochs corresponding to redshifts z between $\approx 10^2$ and ≈ 10. Assuming that the concept of Cold Dark Matter (CDM) provides a correct description of the dark mass, the formation and evolution of the dark-matter halos can be predicted by means of numerical simulations (e.g., [492]). According to the theory the first DM halos were formed in regions with the largest matter overdensity. Since in CDM models a high density contrast is reached first on small scales, the first bound structures had small total masses approximately in the range $10^5 < M/M_\odot < 10^6$. Less massive fluctuations were damped by the Silk mechanism during the radiation-dominated epoch. Moreover, for masses $< 10^5 M_\odot$ the pressure forces were not negligible after recombination. The more massive ($> 10^6 M_\odot$) bound structures which we see today were formed later by merging of the initial low-mass halos.

The DM simulations provide quantitative predictions for various parameters of the dark matter distribution as a function of cosmic time and redshift. Among these parameters are the density profiles, the sizes, the mass spectrum, the correlation length, and the merging rates of the dark halos.

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_10,

More difficult is the prediction of the evolution of the baryonic matter. Just after the epoch of recombination, when the gas pressure effects were still negligible on the relevant scales, the gas density simply followed the dark matter distribution. At these early epochs the dark halos are expected to contain a baryonic mass fraction of the matter similar to the cosmic value (~ 0.18). However, eventually the gas became heated by compression and by the radiation of the first stars and massive black holes. From this point in time the theoretical predictions of the behavior of the gas become less certain. Because of its minor contribution to the total matter density, the gas could not significantly affect the growth of the DM halos. However, the feedback effects from the star formation and stellar evolution could strongly modify the baryonic matter distribution by ejecting gas from the affected halos, or even from neighboring low-mass halos. Moreover, the radiation field of the first stars modified the conditions for star formation in the surrounding gas.

During the past years progress has been made in the theoretical understanding of these processes, and there are major ongoing efforts to further improve the theory (see Sect. 11.3). Qualitatively, it is clear that the gas accumulating in the first low-mass dark halos can reach conditions where star formation can set in (e.g. [254]). It also appears clear that several different mechanisms can lead to the formation of black holes in the centers of the forming galaxies. However, the present hydrodynamical models of the evolution of the gas cannot (yet) cover the range of scales between halos and protostars, and the models do not yet include the full physics of the complex star-formation process. Therefore, some of these processes are approximated by semi-analytical models.

Simulations of the formation of stars and galaxies in dark halos using semi-analytic models do give plausible results on the structure of the resulting stellar systems (e.g., [358]). But, because of approximations and free parameters, it is difficult to estimate the accuracy of the predictions based on the semi-analytic procedures.

In spite of these uncertainties, there are various quantities where a comparison between the theoretical models and the observed parameters of the high-redshift galaxies provide valuable and reliable tests of the theory. Examples are given in the next sections.

10.1.1 Galaxy Sizes as a Function of Redshift

Among the parameters, where a direct comparison is possible, is the size distribution and the size evolution of the galaxies. As pointed out above, galaxies are formed by the infall of baryonic matter in dark-matter halos. According to the semi-analytic models, the angular momentum of the halos and the gas results in the formation of gaseous disks. The models normally assume that stars start to form as soon as the gas surface density in the disks exceeds a certain value [262]. At this point the disks become star-forming galaxies.

N-body simulations of the dark-matter evolution indicate that the angular momentum of the dark-matter halos reach approximately a log-normal distribution of the spin parameter

$$\lambda = J \mid E \mid^{1/2} G^{-1} M_h^{-5/2} \tag{10.1}$$

where J, E, and M_h are, respectively, the angular momentum, the total energy, and the mass of the halo, and G is the gravitational constant. As shown by Shen and colleagues [474], a log-normal distribution of λ results with good approximation in a log-normal distribution of the sizes of the disks formed by the baryonic matter. This theoretical prediction is in good agreement with the observational result (reported in Sects. 2.2.2 and 7.1.2) that the observed distributions of the galaxy sizes are at all redshifts well approximated by a log-normal function.

In Sect. 7.1 it has been shown that many observations indicate an increase of the typical galaxy sizes with decreasing redshift. The models predicts that at high redshift the galaxy sizes scale with the sizes of their corresponding dark halos. As noted in Sect. 3.3.2, at high redshifts, the sizes of halos of constant mass are expected to increase with decreasing redshift approximately $\propto (1+z)^{-1}$ (see Eq. 3.42). According to Bouwens and colleagues [64] for high-redshift starburst galaxies of constant luminosity, the observed sizes are found to vary with the redshift $\propto (1+z)^{-1.1\pm0.3}$. Thus, the observational result are in good agreement with the theoretical predictions of the CDM models.

10.1.2 The Growth of Galaxy Masses

Because of the hierarchical built-up of mass by halo mergers, the theory predicts a growth of the galaxy masses with increasing cosmic time and decreasing redshift (Sect. 3.3.1). Because of the uncertainties in modeling star formation and because of the relatively large error bars of empirically derived masses of high-redshift galaxies, a quantitative statistical comparison is more difficult than in the case of the size evolution. But the evolution of the dark-halo masses derived from N-body simulations provide at least firm upper limits for the expected galaxy masses as a function of redshifts.

Qualitatively, the hierarchical built-up of the masses of galaxies appears to be confirmed by the (relative to the local universe) small typical stellar mass of the observed high-redshift galaxies (e.g., [156]) and by the scarcity of high-mass galaxies observed at high redshifts. On the other hand, there have been reports of individual high-redshift galaxies with unexpectedly high mass values. In some cases additional observations resulted in lower redshifts (e.g., [97, 149]) and luminosities, which removed the apparent contradiction between the theory and the observations in these cases. But there are high-redshift galaxies with unexpectedly high-mass values where these large masses appear to be reliably established. Particularly high masses have been derived for distant submillimeter galaxies. Among these objects is a galaxy with a photometric redshift of $z = 4.7$ and an estimated stellar mass of about $2 \times 10^{12} M_\odot$ [150]. This would correspond to a (minimal) halo mass of the order $10^{13} M_\odot$. According to the dark-matter simulations such massive halos are extremely rare at this redshift (see Fig. 3.12). So far, the mass values of the

high-mass submillimeter galaxies are based on photometric redshifts. Before conclusions from the existence of the high masses at these redshifts are drawn, spectroscopic confirmations of these redshifts and of the nature of these highly interesting objects should be attempted.

In their study of the galaxy masses in the GOODS-MUSIC sample, Elsner et al. [156] found for reliably determined redshifts $z > 4$ only very few galaxies with stellar masses $M > 3 \times 10^{10} M_{\odot}$, and they found no objects with a reliably determined mass $M > 3 \times 10^{11} M_{\odot}$ at these redshifts. Among the high-redshift radio galaxies with well-determined masses there is one object (TN J0924-2201) with a reliable spectroscopic redshift of $z = 5.195$ and an estimated stellar mass of $M \approx 10^{11} M_{\odot}$ [469]. Similar values have been estimated for the dynamical mass of the inner regions of the host galaxies of QSOs with redshifts $z \geq 6$ (e.g., [550]). Since the dark matter is not expected to play a role at the observed small scales, the observed dynamical mass of the high-redshift QSO host galaxies must be of baryonic nature. Obviously, these luminous high-redshift QSO hosts and radio galaxies have for their redshift exceptionally high mass values. However, since the observed space density of these rare objects is very small, their large masses at high redshift can be explained by their formation in rare regions of exceptionally high density. Quantitative estimates confirm that statistically the observed space density of these objects is consistent with the theoretical expectations (e.g., [492]).

10.1.3 The Early Chemical Evolution

The chemical enrichment of the cosmic matter started within a few million years after the beginning of star formation, when the first massive stars exploded as supernovae. Because of the low mass and small gravitational potential of the first dark-matter halos, much of the initially formed heavy elements were probably lost from these halos (and the corresponding galaxies) by supernova-driven galactic winds. The predicted strong mass loss from the first galaxies provides a plausible explanation for the observed early enrichment of the IGM and of the substantial metallicity of the high-redshift IGM gas (Sect. 9.1.2). With increasing mass of the galaxies, mass loss by galactic winds became less important, and an increasing fraction of the heavy elements could be retained by the galaxies.

Many different groups have carried out model computations for the chemical enrichment of galaxies with cosmic time (for references and a thorough discussion of this topic see [330]). These simulations were based either on population synthesis models of single galaxies, or on hydrodynamic calculations of the evolution of the gas and stars in the context of hierarchical evolution models, or on chemodynamical simulations in different dimensions. The predictions of these models can be directly compared to the observational results of the chemical evolution of the high-redshift starburst galaxies, which has been described in Sect. 7.1.2.

All the model computations and the observations agree on the fact that a fast increase of the metallicity took place at high redshifts. But in detail the published

models give different results, and quantitatively most of the theoretical models do not agree well with the observational results [330]. Among the model predictions fitting the (present) observations best are those of A. M. Brooks and colleagues [76].

There are several possible causes for the quantitative disagreement between the models and the observations. Firstly, because of the special conditions in the ISM of the high-redshift starburst galaxies, the calibrations of the observations may contain systematic errors. Secondly, because of the various complex physical processes involved, the chemical enrichment models contain many approximations and simplifications of the relevant physics. More sophisticated future models may result in a better agreement between the theory and the observations (and between the different models).

According to the CDM simulations the most massive present-day galaxies have started to form early in the cosmic regions of exceptionally high matter density. Therefore, in these systems star formation and the chemical enrichment started early, and high star-formation rates resulted in a fast chemical enrichment. In contrast the low-mass galaxies started late and experienced lower SFRs. This obviously provides a natural qualitative explanation of the observed strong dependence of the chemical enrichment on the galaxy mass (Sects. 2.2.3 and 7.1.2 and Fig. 7.21). The effect was possibly enhanced by a dependence of the initial mass function on the environment [273]. Thus, qualitatively, the observed chemical evolution is consistent with the expectations of the CDM models.

10.1.4 Predicted and Observed Space Distribution

As explained in Sect. 8.2, the correlation length of the distribution of the galaxies in the local universe corresponds approximately to that predicted for the distribution of the dark-matter halos at $z = 0$. On the other hand, significantly larger correlation lengths were found for the known high-redshift objects. But, as pointed out in Sect. 8.2, this can be explained by a larger biasing factor of the distant objects. Since the high-redshift galaxies were formed early in dense cosmic regions and since (due to selection effects) the observed samples of high-redshift galaxies contain (relative to the mass function of their epoch) mainly objects of relatively high mass, these high biasing factors are consistent with the predictions of the CDM models. The observed increase of the biasing factors with increasing galaxy mass and with the stellar population age (Fig. 8.3) again is in good agreement with the theoretical predictions and provides further support for the current theoretical concepts.

According to the numerical simulations of the evolution of structure, the most dense and most massive present-day matter concentrations are already discernible in the matter distribution at redshift $z \approx 6$. This is illustrated by Figs. 3.4 and 3.5, where the prominent matter concentration of Figs. 3.5 is clearly visible already at $z = 5.7$. As described in Sect. 8.3 this prediction of the numerical simulations is confirmed by the observational detection of "protoclusters" with redshifts of this order.

The simulations also predict that the first massive black holes were formed and developed in the same very dense matter concentrations [492]. This is confirmed by the observed frequent association of high-redshift protoclusters with luminous radio galaxies and QSOs (Sect. 8.3 and Fig. 8.5). The formation of these structures (only) in regions of exceptionally high matter density also explains why high-redshift protoclusters are very rare objects.

10.2 Constraints on the Beginning of Star Formation

The baryonic mass accumulations in dark halos became galaxies when the first stars started to form in the gas. According to the theory, the conditions for star formation were reached first in the regions of exceptionally high matter density. Star formation became widespread when the gas in the more normal dark halos reached sufficiently high densities. Thus, at different locations star formation started at different times, and the beginning of star formation extended over a substantial time and redshift interval.

In Sect. 2.2.6 the basic mechanisms of the star-formation process have been described. But it has also been noted that many details of the physics of star formation is still not fully understood. As a result, the exact time when the first stars and galaxies occurred cannot yet be predicted reliably from the theory. The current theoretical models indicate that the conditions for first star formation were reached at an epoch corresponding to a redshift between $z = 30$ and $z = 10$. Work which is in progress at present (see Sect. 11.3) may allow us to make more accurate predictions in the near future.

Independently of the gaps in the theory, some firm empirical constraints on the beginning of star formation can be deduced from the observations of the high-redshift galaxies described in Chap. 7. The following are among these firm constraints:

- According to Fig. 7.12, at a redshift of $z = 5$ already about 1% of the present-day stellar mass had been formed.
- Luminous starburst galaxies existed already at redshifts $z > 7$.
- The presence of heavy elements, CO molecules, and dust in the host galaxies of $z > 6$ QSOs shows that in these objects stars have been formed and evolved well before that epoch.

The dust and molecular content of the host of the most distant known QSO indicates that about 10^8 SNe must have exploded in this galaxy before the observed light was emitted. With a normal IMF this corresponds to an integrated star formation of about $10^{10} M_\odot$. Assuming an average star-formation rate of $10^2 M_\odot$ yr^{-1} this would indicate a beginning of the star formation in these objects about 10^8 years before the observed light was emitted about 9×10^8 years after the big bang.

Since we have no direct knowledge on the SFR and the IMF at these early epochs, these estimates cannot provide reliable constraints on the epoch of the first star

formation. However, the numbers show that the observations can be explained with reasonable assumptions.

Somewhat more conclusive estimates are possible using galaxies where star-formation ages can be estimated by means of well-determined spectral energy distributions. Most galaxies, for which adequate photometric data exist, have relatively low redshifts and, therefore, do not provide information on populations originating at redshifts $z > 6$. An exception is the relatively well-studied lensed galaxy A1689-zD1, which has been investigated by Bradley and colleagues [68]. According to these authors A1689-zD1 has a redshift of $z \approx 7.6$, a stellar mass between $1.6 \times 10^9 M_\odot$ and $3.9 \times 10^9 M_\odot$, and a stellar population age between 45 Myr and 320 Myr. Thus, in this galaxy significant star formation seems to have started at a cosmic age between 0.36 Gyr and 0.63 Gyr (corresponding to $11 \geq z \geq 8$).

Using less accurate photometric data of a sample of 16 galaxies with likely redshifts $z \approx 6$, formation redshifts in the range $7 < z < 18$ have been estimated by Eyles et al. [161]. More accurate photometric observations covering a larger wavelength range of this galaxy sample will hopefully allow us to constrain the age of the stars in these galaxies more precisely.

From the observations quoted above it seems well established that star formation started at an epoch corresponding to a redshift $z > 8$ (i.e., at a cosmic age < 0.6 Gyr). Again this observational result is consistent with the theoretical predictions. But a more precise derivation of the epoch of the first stars is obviously highly desirable.

10.3 Properties of the First Stars

As pointed out in Sect. 3.3.2, the first stars initially were composed of a pure hydrogen–helium gas. The current theory, discussed in Sect. 3.3.2, predicts that these first-generation "Population III" stars were massive, hot, and shortlived. They exploded as SNe after a lifetime of the order 10^6 years. The theory also predicts that no Population III stars with masses less than a few solar masses were formed. Since only low-mass stars (with approximately $M < M_\odot$) would have evolved slowly enough to be still observable since the epoch of the first star formation, Population III stars should not be detectable in the local universe. This seems to be confirmed by the fact that, in spite of extensive searches, no low-mass stars with a pure hydrogen–helium composition could be found in the Milky Way galaxy.

Theoretical models of Population III stars have been computed by various different authors (see, e.g., [452] and the literature cited there). These models predict for an initial pure hydrogen–helium composition effective temperatures of zero-age main-sequence stars up to about 10^5 K and luminosities up to about $10^7 L_\odot$. A heavy-element abundance as low as 10^{-7} reduces the effective temperature significantly, while the luminosity is much less affected. Normal luminous hot stars are characterized by strong stellar winds, which result from the absorption of the momentum of the photospheric radiation by ions such as C IV, O VI, Si IV, etc.

Since such ions are absent in Population III stars, these stars are not expected to have strong stellar winds.

According to the present theoretical predictions, the Population III stars were formed singly or in small groups in low-mass halos at redshifts $z > 10$. Assuming that the light of these objects could escape their environment, they could in principle be observed as faint infrared sources. Detecting these individual stars or small groups is outside our present observational capabilities. However, it has been suggested that the unresolved, redshifted radiation of these objects could be the origin of unexplained fluctuations of the IR background in observations obtained with the Spitzer IR satellite [260, 261]. But there exist other possible mechanisms which could cause the observed fluctuations of the infrared background. Therefore, it is not clear at present, to what extent light from Population III stars contributes to this background radiation.

Because of the high effective temperature and the absence of strong radiation-driven stellar winds massive Population III main-sequence stars produce a large amount of FUV and EUV photons. As a result, such stars are surrounded by hot H II regions with characteristic spectral properties. Theoretical predictions for the resulting spectra of galaxies containing Population III stars have been published, e.g., in [452] and [530]. Among the main properties of these predicted spectra are extremely large Lyα equivalent widths (up to 1500 Å). Because of the high effective temperature of the Population III stars, He II recombination lines are also expected from the ionized interstellar gas of such galaxies. For the He II 1641 Å line equivalent widths between about 20 Å and 80 Å have been predicted for a very young pure Population III.

As pointed out in Sect. 7.1.3 the observed high-redshift galaxies have, with few exceptions, Lyα equivalent widths $<$250 Å. These equivalent widths can be explained without invoking Population III stars. The few observed larger equivalent widths can be explained by assuming for these objects an exceptionally young Population II.

He II 1641 Å emission is observed in part of the high-redshift galaxies (see, e.g., Fig. 5.1). However, so far, the equivalent widths of the He II emission lines in high S/N spectra of high-redshift starburst galaxies have been found to be always $\leq$2 Å. This is significantly smaller than the values predicted for the H II regions of the Population III stars. In all cases the observed He II emission can be explained by the presence of Wolf-Rayet stars in a very young Population II (e.g., [71]). Dedicated searches among the high-redshift galaxies with redshifts $z \approx 4$ and $z \approx 4.7$ also did not turn up He II emitter which could be convincingly ascribed to Population III stars [368].

Another way to trace Population III stars is to look for the chemical signature of the supernovae produced by these objects. As pointed out in Sect. 2.1, very massive stars with a low heavy-element content are expected to produce pair-instability supernovae (PISNe), which produce distinct chemical abundance patterns. No such signatures have been found in low-metallicity galactic stars. However, this can possibly be explained by an "overshooting" of the chemical enrichment by the massive Population III star explosions. In this case the chemical abundance pattern of PISNe

is expected to occur in stars with a moderate chemical abundances, where this signature would be diluted by the effects of Population II stars [257].

Even at stellar masses well below those required for pair-instability supernovae, exploding Population III stars may result in abundances of the ejected material which is distinctly different from that of Population II SNe [533]. These predicted abundance patterns have been compared to the chemical abundances of the most metal-poor (and presumably oldest) stars observed in the Milky Way galaxy, the relative abundances observed in high-redshift galaxies, and with the relative abundances found in the IGM. In an important paper Umeda and Nomoto [533] could prove that Population III supernova models can indeed reproduce the observed unusual relative abundances of galactic stars with iron-to-hydrogen ratios lower than 10^{-5} of the solar value. However, subsequent work has shown that these abundance patterns (characterized by a huge overabundance of carbon) can also be explained by secondary effects in objects with a small, but non-zero initial metallicity and an initial more normal abundance pattern (e.g., [36, 531]).

As noted in Sect. 7.1.2, the relative abundances of the ISM and the stars of high-redshift galaxies are generally in agreement with the predictions of Population II nucleosynthesis. The same is true for the IGM abundances. However, at present, abundance derivations for the distant galaxies and IGM filaments are much less accurate and less complete than those of the old galactic stars. Future observations of higher quality may well turn up deviations from the Population II predictions.

Population III supernovae are also expected to produce copious amounts of dust [380]. This dust again shows distinct properties, which in principle could be used to identify its origin from Population III SNe. So far, no attempt seems to have been made to identify the dust from Population III SNe observationally.

Summarizing, we have to note that the observational programs to identify the predicted Population III stars or their remains or effects have not been successful so far. In spite of this negative result, it is clear from the theory that Population III stars once existed and that they played an important role in the cosmic evolution. There are two possible explanations for the fact that no conclusive observational evidence for these objects could be found: Either the epoch of Population III stars was very short, or the properties of these objects were different from the current predictions.

10.4 The Passively Evolving High-Redshift Galaxies

The first galaxies were—by definition—star-forming systems. Passively evolving systems were formed when some of the star-forming galaxies lost all their cool gas and stopped forming new stars. As noted earlier, there are several possible mechanisms which can remove the cool gas from a galaxy. One possibility is that the gas is simply exhausted by intense star-formation activity. But the gas can also be removed by dynamical interactions or a merger with another galaxy. Finally, the gas can be blown out of the galaxy by stellar winds, by SN outflows, or by the outflows and radiation pressure of a central AGN.

Observational evidence suggests that gas loss due to AGN activity plays a important role for the formation of passively evolving galaxies at low redshifts (e.g., [454]). The effects of AGN for high-redshift galaxies is less clear, since the mechanisms and the epoch of the formation of the central black holes in galaxies are not sufficiently well known at present (Sect. 10.5).

Some information on the potential role of the AGN on the formation of passively evolving galaxies can be deduced from the observed history of the QSOs and their host galaxies. Observations of the high-redshift QSOs show that black holes with masses $>10^9 M_\odot$ existed already at redshifts $z = 6.4$. As pointed out in Sect. 7.3, most of these distant QSO hosts galaxies with redshifts $z > 3$ show evidence for intense star formation and for the presence of cool interstellar matter. Even at redshifts $2 < z < 3$ QSO host galaxies often are found to contain young stellar populations. This seems to indicate that in massive high-redshift galaxies star formation persists (at least for some time) in spite of intense AGN activity, and that the cool gas was not removed immediately. Alternatively, if gas is removed by the AGN activity, it is replenished by mergers with gas-rich systems.

In the presently available samples of high-redshift galaxies, significant numbers of passively evolving "red-sequence" galaxies are detectable only at redshifts $z \leq 3$ (see Sect. 7.2). Red galaxies with predominant passively evolving populations at $2 < z < 3$ often show evidence for residual star formation activity or the presence of intermediate age stars (Sect. 7.2). Assuming that this results are not due to selection effects, these findings indicate that passively evolving galaxies are found only at epochs corresponding to $z \leq 3$ and that the redshift range $2 < z < 3$ corresponds to the main formation period of these systems.

As noted in Sect. 7.2, at all redshifts the passively evolving galaxies are on average more massive than the starburst galaxies observed at the same redshift. Obviously, most of the star formation in these massive systems must have taken place at earlier epochs. This dependence of the star-formation activity on the mass is even more pronounced in the local universe where practically all very massive galaxies are passively evolving ellipticals (Sect. 2.2.2). In contrast, as noted above, at $z > 5$ we observe intense star formation in the (for that epoch) most massive galaxies. Obviously, with increasing cosmic time and decreasing redshift intense star-formation activity moves to galaxies with progressively lower mass. As pointed out in Sect. 7.1.2, in the astronomical literature this effect is called "downsizing."

The observational facts summarized above and the constraints of the CDM theory lead to the following consistent scenario of the formation of the passively evolving high-redshift galaxies.

According to the cold dark matter simulations the massive galaxies started to form early in dense regions of the universe. Star formation and stellar evolution started early in these systems, resulting in the presence of an evolved stellar population already at high redshift. The early evolution of these galaxies took place in a high-density environment. Therefore, these galaxies very likely experienced early multiple mergers with (other) gas-rich galaxies. In these mergers part of interstellar gas was probably lost. Very likely, strong starbursts and (with a slight delay) AGN activity was triggered by these mergers. These processes probably resulted in

a further loss of the interstellar matter. Eventually this resulted in a termination of star formation and the emergence of high-redshift passively evolving systems. Since with increasing cosmic age the merger rates were decreasing, a replenishment of the gas by the capture of gas-rich galaxies became less likely.

The central AGNs probably contributed to the loss of the interstellar gas. But at high redshift this mechanisms appears to have been less efficient and less important than in the low-redshift universe.

10.5 Formation of the Supermassive Black Holes

Since CDM is dissipationless, it cannot cool to form black holes, and hence the formation of these objects cannot be studied by CDM simulations. Hydrodynamic simulations of the cosmic structure formation usually do not have enough resolution to model the black holes. Therefore, in simulations of the formation of galaxies the supermassive black holes in the center of these objects generally are included by means of semi-analytic approximations. This is further complicated by the fact that there exist several different physical scenarios for the formation of these objects.

Some authors have suggested that the massive black holes in the cores of present-day galaxies resulted from rapid mass accretion and merging of stellar-mass "feed black holes" which were formed in SN explosions of the first very massive (Population III) stars (e.g., [328]).

On the other hand, massive black holes may also have been formed directly by the accumulation of gas in dark-matter halos, either from gas with exceptionally low angular momentum ([154]), or by dynamical instabilities of rotating gas masses (e.g., [37]), or by the inhibition of fragmentation in the collapsing gas of early dark halos ([490]).

Independently on how the black hole initially forms, the computations show that a black hole in the center of a dark halo grows rapidly by further mass accretion and eventually becomes a supermassive black hole, which (as observed) scales with the bulge mass of the final galaxy (e.g., [137]).

The different black-hole formation mechanisms predict different formation epochs. But none of the processes is understood well enough to provide accurate quantitative predictions. Observations of the host galaxies of the most distant QSOs show that in these systems a large fraction of the baryonic mass is still present in the form of gas, while the ratio between the stellar mass and the black hole mass is much smaller than in the local universe. This may indicate that a direct formation of the black holes from primordial gas is the most likely initial black-hole formation process.

10.6 The Progenitors of the Present-Day Galaxies

Because of mergers, mass accretion, the loss of gas, and the evolution of the stellar population, the galaxies changed their properties significantly between the epochs corresponding to $z > 2$ and $z = 0$. Computer simulations of the evolution of

galaxies and a comparison of the observed properties of the high-redshift galaxies with locally observed stellar systems make it possible to relate the present-day galaxies with their high-redshift progenitors (e.g., [300]).

The most important parameters influencing the evolution of galaxies is the star-formation rate. At least at low and intermediate redshifts the SFR of star-forming galaxies depends on the mass of the stellar systems. According to K. Noeske and colleagues star-forming galaxies are located along a tight sequence of SFR vs. stellar mass, with the logarithm of the SFR being approximately proportional to the logarithm of the stellar mass [376]. Therefore, massive star-forming galaxies are expected to evolve faster than their low-mass counterparts.

Among the most massive star-forming galaxies at high redshift are the hosts of the most luminous AGN observed at $z > 5$ (Sect. 7.3). All current evolutionary models suggest that these galaxies have been the progenitors of the central galaxies of the most massive present-day galaxy clusters (e.g., [492]). These central cluster objects typically are giant elliptical galaxies or cD systems. To accumulate the observed large mass, these galaxies must have formed early in regions of high cosmic matter density. Star formation started (and probably ended) early in these systems. The protoclusters observed at high-redshift appear to be the precursors of the bound massive clusters which we see today. Detailed observations of the population content of the central galaxies of massive clusters would be of great interest, since they may allow us to derive the exact epoch, when the star formation took place and ended in these massive systems. Although derived from local observations, these data are expected to provide valuable information on the high-redshift universe.

The observed luminous high-redshift submillimeter galaxies with redshifts $z > 4$ and extreme star-formation rates ($> 10^3 M_\odot$ yr^{-1}) are the natural candidates for the progenitors of the passively evolving massive galaxies observed at $z \geq 2$ (e.g., [87, 99]). The high SFRs of the luminous submm galaxies almost certainly result from massive mergers. Because of their extreme star-formation rates most of their interstellar matter is depleted on a very short timescale. As a result, these systems naturally evolve into dense galaxies with a high stellar mass and a low star-formation rate.

Those luminous dust-obscured starburst galaxies observed at redshifts $z \sim 2$, which have been investigated in detail, were found to be massive and dense systems. At least in some cases, these galaxies consume their gas on a timescale comparable to the local dynamical timescale. Very likely these galaxies will end up as massive elliptical galaxies of the local universe (e.g., [324, 505]).

While this explains the origin of some of the present-day elliptical galaxies, the origin of the majority of the present-day elliptical galaxies are less obvious. The basic formation mechanism of the passively evolving galaxies has been outlined in Sect. 10.4. As pointed out there, since there are different mechanisms which can result in the removal of the gas from these systems, there still exist uncertainties concerning the evolution of these systems. But, as noted in Sect. 7.2, the high central matter densities of the known high-redshift passively evolving systems and the low number density of the passively evolving systems at $z = 2$ seem to rule out that these systems are the direct progenitors of the typical present-day elliptical galaxies and

the bulges of spirals. While part of the elliptical galaxies may have been formed by an AGN-driven loss of gas (see, e.g., [454]), others may have been formed (and lost their gas) by mergers. In view of the age of the stellar populations of the elliptical galaxies, these mergers must have taken place long ago, or there must have been "dry mergers" without further star formation. Observational evidence that many of the present-day massive elliptical galaxies underwent major mergers since $z \approx 1$ has been described, e.g., in [40], [41], and [316]. However, other studies find dry-merger rates which are too low to impact the evolution of these objects significantly [451].

There is better agreement on the evolution of the (less massive) high-redshift UV-bright starburst galaxies. The space density of the typical UV-bright starburst galaxies at $z = 3$ is similar to the local density of L^* galaxies. Due to the rapid exhaustion of the cool gas, the star-formation rates of these objects had to decrease with cosmic time. A majority of the UV-bright starburst galaxies, including the Lyman-break galaxies observed at $z = 3$, have probably evolved into present-day mildly star-forming $\sim L^*$ systems, such as our Milky Way galaxy (e.g., [110, 372, 501]). The stars formed in the high-redshift starbursts are expected to form the bulge population of the present-day galaxies. Some of the high-redshift starburst galaxies may have lost all their cool gas through interactions. In these cases the former starburst galaxies evolved into low-luminosity passively evolving systems of the present-day universe. The low-mass high-redshift starburst galaxies observed as Lyα emitters probably evolved into low-luminosity (still) metal-poor dwarf galaxies.

Most local star-forming galaxies have properties which are quite different from the typical high-redshift starburst galaxies. The present-day galaxies typically are more massive, less luminous, less concentrated, and more enriched with heavy elements. However, there is some overlap of the observed properties of the very luminous high-redshift dust-obscured starburst galaxies and the local LIRGs and ULIRGs. Obviously the LIRGs and ULIRGs reached a similar evolutionary state as their high-redshift counterparts at a later cosmic time. Moreover, the low-mass local star-forming galaxies known as "blue compact dwarfs" (BCDs) have properties similar to those of low-mass Lyman-break galaxies. Very likely, these low-mass objects are evolving slowly and have been lagging behind compared to the more massive objects. Like most high-redshift galaxies the BCDs have low metallicities. But their SFRs typically are smaller than in LBGs of the same mass.

As noted first by T. Heckman and colleagues [224], there exists a rare population of local starburst galaxies with properties overlapping those observed in UV-bright galaxies at $z \approx 3$. They have low metallicities, stellar masses of the order $10^{10} M_{\odot}$, and most of their stars seem to have been formed during the past 1–2 Gyr. Since these galaxies are nearby objects, they can be studied in much more detail than their high-redshift counterparts.

Figure 10.1 shows examples of HST images of such local ($0.09 < z < 0.17$) analogs of distant starburst galaxies. As shown by the figure, all observed objects show evidence for mergers or interactions. At least in one case (SDSS J032845.99+011150.8) tidal tails are clearly visible. The observations strongly support the hypothesis that mergers and interactions are the main cause of starburst activity in galaxies.Very likely these "local LBGs" have formed in the same way as

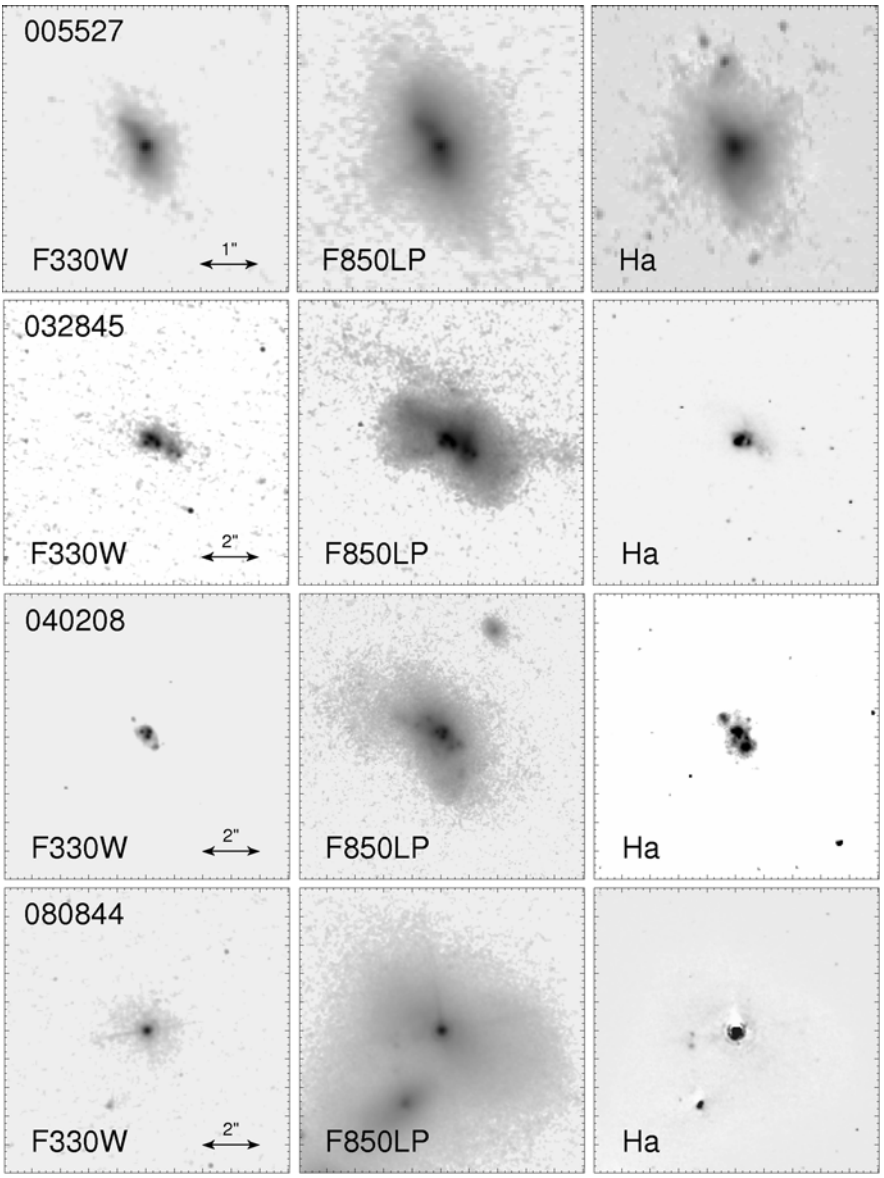

Fig. 10.1 HST images of local starburst galaxies which have properties similar to the LBGs observed at high redshift. For each galaxy images in the UV (330 nm), the red spectral range (850 nm) and Hα are given. The individual galaxies are identified by the first six digits of their SDSS designation. Detailed information on the individual objects can be found in [398]. (From Overzier et al. [398])

their high-redshift counterparts, i.e. through recent mergers. However, in the local objects these mergers took place at a much later cosmic epoch. As pointed out by Heckman and colleagues, in these objects we observe “living fossils” of the star-formation history.

Chapter 11
Ongoing Work

In the preface to this book it has been noted that the high-redshift galaxies form a topic where much work is still in progress. Some of the current research projects in this field promise significant advances in our understanding of the high-redshift universe. This chapter will describe some examples.

Part of the ongoing large-scale projects (such as the COSMOS and GOODS projects) have already produced significant new data. Therefore, these programs have already been introduced in Sects. 5.2 and 7.1.1. In the following, some promising new programs will be listed, which have started only recently. In addition, some observational programs will be mentioned which are still in a preparation phase, but where results are expected to become available in the very near future.

11.1 Enlarging the Data Base

Part of the recent and near-future programs aim at extending the available data base of high-redshift objects by means of new large surveys. The objective of these projects is to improve our knowledge by generating statistically more significant samples of galaxies of different types. As pointed out in Sect. 5.2 programs aiming at observing a large number of objects or sampling large volumes of space are most efficiently carried out using dedicated telescopes. A notable and highly successful example (described in Sect. 5.2) has been the Sloan Digital Sky Survey (SDSS), which has been carried out with a dedicated 2.5-m telescope at the Apache Point Observatory in New Mexico. In addition to the original SDSS I, an SDSS II has recently been completed. The SDSS II applied the techniques of the original SDSS to a smaller equatorial stripe of the sky, which was observed repeatedly to find variable objects and to reach deeper. Some important results of the SDSS II have already been mentioned in Sect. 9.2.1. A third SDSS survey is in preparation at present. The extragalactic part of the SDSS III has been named the "Baryon Oscillation Spectroscopic Survey" (BOSS). In the framework of the BOSS survey spectra of $\sim 1.5 \times 10^6$ galaxies with redshifts <0.7 and spectra of the Lyman forest of about 160 000 QSOs in the redshift range 2.2 to 3 will be taken. As indicated by the program name, the main objective of BOSS will be to measure the baryon oscillations (see Sect. 3.2.3)

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_11,

at different redshifts. However, the program will also yield information on other aspects of the extragalactic universe. Part of these additional data will be of interest for the field of high-redshift galaxies. According to present plans, the SDSS III will start in 2009, and first data will be released in 2010. (For more information see: www.sdss3.org.)

There are several other ongoing or planned programs aiming at measuring the baryonic oscillations in the space distribution of galaxies. A large program of this type in the southern hemisphere is the "WiggleZ" project [206], which is being carried out at the 3.9-m Anglo-Australian Telescope of the Siding Spring Observatory. WiggleZ will result in spectra of 400 000 star-forming galaxies with redshifts $z \leq 1$. The field to be surveyed has a size of about 10^3 square degrees.

Another major new survey program using dedicated telescopes is the "Panoramic Survey Telescope and Rapid Response System" (Pan-STARRS) project. This program is being carried out by the University of Hawaii, with participations of universities and research institutes in the US, Europe, and Asia. Pan-STARRS uses four 1.8-m telescopes, each having a 3-degree field of view. Each telescope will be equipped with a 1.4×10^9 pixel camera for the visual wavelength bands. The system will image the part of the sky which is observable at Hawaii with multiple short exposures. After a full image of the observable sky is obtained in about one week, a new cycle of sky exposures will be started. The main objective of the project is to find time-variable and moving objects. However, by stacking all exposures of a given area of the sky, deep sky maps will be obtained after some years of operation. These deep sky maps will form a valuable data base for studies of the high-redshift universe.

So far, one of the planned four Pan-STARRS 1.8-m telescope has become operational. At present this telescope is used mainly for developing the system. The timescale for the full survey program and for the expected release of the deep stacked images, which will contain information on high-redshift galaxies, is still uncertain at present. For more information see pan-starrs.ifa.hawaii.edu.

A larger (4-m aperture) new dedicated survey telescope is the "Visible and Infrared Survey Telescope for Astronomy" (VISTA), which has been developed in the UK and which will be operated at the ESO Paranal Observatory in Chile. The VISTA telescope is devoted to deep surveys in the NIR. The telescope has a 1.6 degree field of view and a NIR camera with 6.7×10^7 pixels. The wavelength range of the camera is 0.88 μm (z band) to 2.15 μm (K band). Several large survey programs are planned with this telescope. Details and status information can be found on the ESO web page (www.eso.org). Most interesting among the planned VISTA programs is the planned "UltraVISTA" survey, which aims at ultra-deep NIR photometry ($K \leq 25.6$) of a 0.72 square degree field, which has been extensively studied already in the course of the COSMOS survey. Among the objectives of UltraVISTA are the detection of LAEs at $z \approx 8.8$ and starburst galaxies at $z \leq 9$.

The assembly of the VISTA telescope has been essentially completed in the first half of 2008 (see Fig. 11.1). The first observations are expected to begin in the second half of 2009.

Fig. 11.1 The 4-m VISTA survey telescope at the Paranal Observatory during the installation of the primary mirror. (For inserting the primary the secondary mirror assembly has been removed). Note the very compact structure of this large wide-field telescope. (Photo European Southern Observatory)

A second dedicated survey telescope at the ESO Paranal Observatory is the "VLT Survey Telescope" (VST). The VST has a 2.6-m aperture and a one square degree field of view. It will be equipped with a 16k $\times$ 16k CCD camera and it will be dedicated to do deep surveys in the optical bands. While the camera (known as "Omega-CAM") has been completed, the telescope is still under construction. Information on the currently planned VST surveys can again be found on the ESO web site.

A deep NIR survey program using an 8-m telescope is the "Dark Ages z Lyman-α Explorer" (DAzLE) project [234]. DAzLE uses a specially designed NIR narrow-band differential imager at the UT3 telescope of the ESO VLT to search for Lyα emitting galaxies at redshifts $z > 7$. The field of view of the DAzLE camera is 6.8 arcmin $\times$ 6.8 arcmin. The first observations with DAzLE were carried out in November 2006 [343].

Another recently started NIR survey is the "NEWFIRM Medium-Band Survey" [539]. This survey makes use of the "NOAO Extremely Wide-Field Infrared Imager" (NEWFIRM) at the 4.1-m telescope of the Kitt Peak Observatory in Arizona. Using a set of six medium-band NIR filters, the NEWFIRM survey will obtain photometric information and photometric redshifts of IR-bright galaxies in the redshift range $1.5 < z < 3.5$. The observations are being carried out in the COSMOS and AEGIS fields. In both cases the field size will correspond to the 28 arcmin $\times$ 28 arcmin field of view of the NEWFIRM camera. This survey is expected to provide for the first time a large sample of red high-redshift galaxies with accurate ($\Delta(1 + z) < 0.03$) photometric redshifts. The survey has been started in March 2008.

A new survey instrument with a particularly large field of view is the 4-m "Large Sky Area Multi-Object Fibre Spectroscopic Telescope" (LAMOST, see

Fig. 11.2 The 4-m LAMOST Schmidt-type wide-field telescope. The primary mirror is housed in the dome to the right. The primary sends the light to a large stationary reflecting Schmidt corrector, which is located at the upper end of the inclined enclosure to the left. The focal plane (with a 5 degree field of view) is near the lower end of the inclined enclosure. (Photo Courtesy Norbert Christlieb)

Fig. 11.2). This telescope has recently begun partial operations at the Xinlong Observatory in Northern China. LAMOST is a Schmidt-type, wide-field telescope with a 4-m primary mirror and a large reflecting corrector. The telescope has a 5 degree field of view, and it is equipped with large multi-fiber spectrographs (see www.lamost.org/en/). When complete, the instrument will be able to produce up to 4000 spectra simultaneously. Part of the observing time will be devoted to obtain spectra of $\sim 10^7$ galaxies. With its 4-m effective aperture and the relatively high background of the fiber technique, LAMOST will not reach as deep as the present $\geq$8-m telescopes. But due to its large field it will be capable of observing efficiently large volumes. Therefore, LAMOST has significant potential for identifying and studying rare intrinsically bright high-redshift objects.

11.2 Efforts to Derive More Accurate Physical Parameters

In addition to the observing programs extending the samples of high-redshift galaxies, there are also important ongoing projects which aim at deriving more accurate and more detailed physical parameters of these systems.

Particularly important for our understanding of the high-redshift galaxies will be improved and more reliable information on the geometric structure of these objects.

With the present-day large telescopes and AO systems galaxies can be resolved at practically any redshift. However, many of the published images of the high-redshift galaxies contain only a few resolution elements. Because of the decrease of the physical size of galaxies with increasing redshift, and because of the rapidly decreasing surface rightness of the very distant galaxies, very little is known of the morphology of the interesting galaxies observed at the highest redshifts. To test the models of galaxy formation it will be important to get better resolved and higher-quality images of these very distant galaxies. The James Webb Space Telescope will give better S/N images of the distant galaxies. But, due to its moderate aperture, there will be no gain in resolution relative to the large ground-based telescopes with AO systems.

Among the already existing facilities the "Large Binocular Telescope" (LBT, see medusa.as.arizona.edu/lbto) has the greatest potential for improving our knowledge on the spatial structure of the high-redshift galaxies. The LBT (Fig. 11.3) consists of two 8.4-m mirrors on a common mount. It has the light collecting power of a $\approx$ 12-m telescope and (when the beams from the two mirrors are combined coherently) the angular resolution of a telescope with a 23-m aperture. With the two mirrors on the same mount, the beams can be combined to form a Fizeau interferometer.

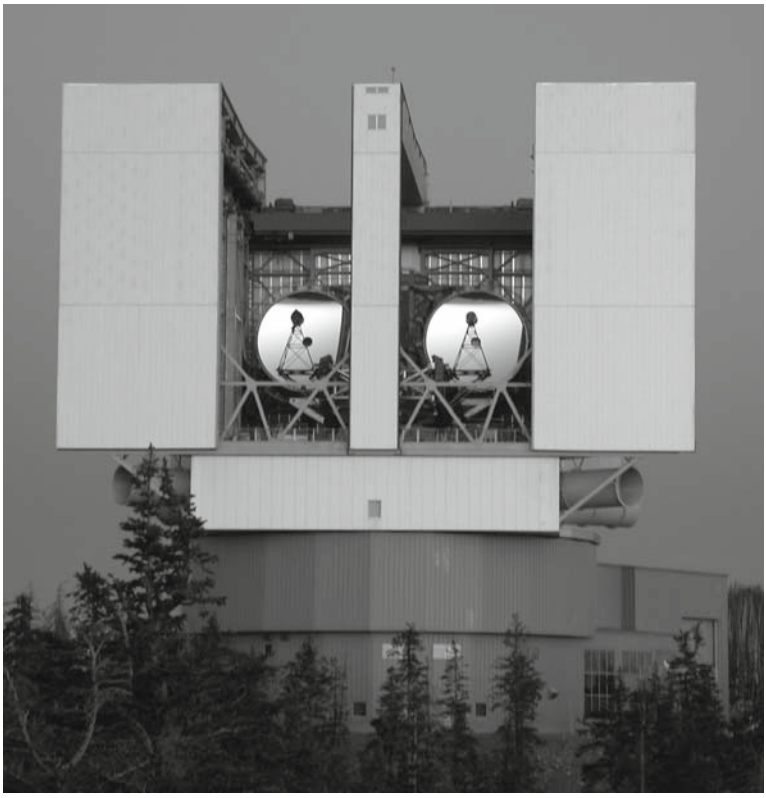

Fig. 11.3 The Large Binocular Telescope. (Courtesy LBT Observatory)

In this way it will be possible to produce directly high-resolution images in a 10–20 arcsec field. The angular resolution is expected to be about 0.02 arcsec in the K band, corresponding to about 0.16 kpc at a redshift $z = 2$.

The LBT has been in limited operation since about 2006 and has already produced valuable scientific results. However, so far only seeing-limited observations have been possible. A Fizeau interferometer is under construction (see www.mpia.de/LINC/). The timeline for the commissioning of the interferometer on the telescope will depend on the installation of the adaptive secondaries of the LBT, which are under construction at present.

Important new information on the physical properties of the high-redshift galaxies are also expected from observations with new powerful infrared spectrographs at the existing $\geq$8-m telescopes. Several such new instruments are presently under construction or have been commissioned recently. Compared to most previous instruments, the new IR spectrographs have larger fields of view, efficient multi-object spectroscopy capabilities, and detectors with much improved noise properties. This will make it possible to get good S/N IR spectra of faint objects in the dark wavelength intervals between the atmospheric OH lines. Observing programs, which are in preparation and which will be carried out with these new spectroscopic instruments, are expected to result in much more reliable constraints on the stellar populations and the stellar ages of the high-redshift galaxies of all types.

11.3 Theoretical Work

In parallel to new observing programs, major current efforts are being undertaken to improve our theoretical understanding of the available and future observational results. Compared to the big observational programs, the theoretical work is carried out by smaller groups, which are distributed over many different institutes. Because of the wide distribution of the work and the diversity of the detailed topics, only a few important examples can be mentioned in the following.

11.3.1 Improving the Spectral Synthesis Models

A critical interface between the observations and the theory of the high-redshift galaxies are the synthetic spectra. Much of the results in terms of physical parameters depend directly on the comparison of observed spectra or colors with those predicted by spectral synthesis models.

During the past years improvements of the spectral libraries have been made by including stellar models with more complete representation of physical effects. An important addition to the published libraries has been the inclusion of rotational effects in the evolution of massive stars (e.g., [309]). Moreover, the metallicity range of the libraries has been extended. However, this metallicity range is still limited.

New efforts to extend the libraries to even lower metallicities will be particularly valuable.

Even more important is the (correct) inclusion of stellar types and evolutionary stages which have been missing so far, or whose treatment has been controversial in the past. A notable example are the thermally pulsing asymptotic giant branch (TP-AGB) stars, which contribute to the NIR spectra of star-forming galaxies. The effects of the contributions of the TP-AGB stars to the spectra of high-redshift galaxies has been studied, e.g., by C. Maraston and colleagues [337], who found that TP-AGB stars strongly influence the spectral energy distribution of the high-redshift galaxies. Important new contributions to this issue are new observations of the NIR SEDs of local or low-redshift galaxies with different known physical parameters (e.g., [157]).

As pointed out in Sect. 7.1.2 a comparison with synthetic spectra published by different groups sometimes leads to different quantitative conclusions. Obviously, at least some of, these models require improvements. Moreover, all population synthesis models contain assumptions which are difficult to verify.

Prominent among these assumptions is the initial mass function. There is no doubt about the importance of this parameter for the interpretation of the synthetic spectra (e.g., [537]). Evidence for a variation of the IMF with cosmic time has been reported, e.g., in [296]. There is general agreement that the IMF of the very first (Population III) stars was significantly different from that observed in the present-day Milky Way galaxy. On the other hand, for all later epochs normally a Salpeter-type (or similar) IMF (see Sect. 2.1.2) is assumed. Although observations in the local universe indicate a surprisingly uniform IMF, there is little direct evidence for the IMF in high-redshift galaxies. Since certain parameters of the distant galaxies, such as metallicity, mass density, gas temperature, and star formation rates, are known to differ from those of local systems, it seems justified to suspect that the IMF could be different in these systems. Ongoing theoretical studies are looking into the dependence of the IMF on the parameters of the galaxies and on the properties of the star formation environment. The results of this work will make the comparison of high-redshift observations with synthetic spectra more reliable.

An interesting example of such work is a study by R. Kleesen and colleagues of the fragmentation process under the conditions in the dense and warm central region of our Milky Way galaxy [269]. The computations of these authors predict for such regions a top-heavy IMF. According to this work, practically no stars less massive than 7 $M_{\odot}$ are expected to form in this environment. As pointed out in [269], this result may have significant implications on the analysis of high-redshift galaxies.

11.3.2 Physical Processes at the End of the Dark Age

The excellent agreement between the observed properties of the cosmic microwave background and the theoretical predictions of CDM models (e.g., Fig. 3.8) seems to show that the physics of the universe at the time of recombination is reasonably

well understood by now. One reason for the good agreement is the applicability of a linear theory for the evolution of structure at this epoch. During the following epochs nonlinear effects become important and the physical processes become more complex. Moreover, at present, there are no direct observations of the epochs between recombination and the epoch at which the most distant known galaxies at redshifts ≈ 8 are observed. As pointed out in Sect. 12.1.1, future radio telescopes may be able to observe the neutral hydrogen gas between redshifts $z \approx 200$ and the epoch of reionization. But the present instruments are not sensitive enough for such observations. As a result, our knowledge on this epoch, which defines the initial conditions of high-redshift galaxies, depends on theoretical models. Obviously, improving these theoretical models is important for our understanding of the high-redshift galaxies.

Among the key issues is the formation of the first stars. As discussed in Sect. 3.3.2, there is a general agreement that the accumulation of gas in low-mass dark-matter halos resulted in dense and cool neutral gas condensations, where massive stars could form, starting at redshifts $30 < z < 10$. As pointed out in Chap. 3, these first stars were formed from gas consisting of hydrogen and helium only. The main cooling agents during the formation of these "Population III" stars were H_2 and HD molecules. The molecular cooling also determined the mass spectrum of the first stars. Among the uncertainties of the present theory are the feedback effects of the first stars on the IGM and the formation of the next generation of stars. As pointed out in Sects. 3.3.2 and 9.2.1, the main effect of the first stars was the heating and ionization of their surroundings by UV radiation. This radiation can lead to the destruction of the cooling molecules, which would inhibit further star formation. However, after the first stars have ended their short lives, the free electrons in their relic H II regions can facilitate the formation of H_2 and HD molecules and, in this way, favor further star formation. (For a detailed discussion of this issue see, e.g., the recent review by J. Johnson et al. [254].) Current comprehensive hydrodynamic simulations which include these feedback processes (see, e.g., [209, 255]) will hopefully result in reliable predictions on the timeline, the stellar IMF, and the galaxy properties resulting from the beginning of star formation.

Another potentially important feedback mechanism is the formation of the first massive black holes in the centers of galaxies. Qualitatively, the radiation from the accretion disks of the first black holes had a similar effect as the first hot stars.

At present, the theoretical predictions on the formation of the first black holes are even less clear than in the case of the first stars. As pointed out in Sect. 10.5, there are several different, equally plausible scenarios for the formation of the supermassive black holes in the centers of the galaxies. These different mechanisms result in different formation times and different feedback effects. Thus, a reliable identification of the formation mechanism(s) of the massive black holes in the galaxy centers is one of the most important issues of the theory of the baryonic matter at the end of the dark age. Improved hydrodynamic computations modeling the black-hole formation are being carried out presently by several different groups. They may clarify the mechanism and the epoch of the formation of the massive black holes in the centers of the high-redshift galaxies.

Chapter 12
Future Facilities and Their Opportunities

12.1 New Instrumentation

Because of the large distance of astronomical objects, astronomers are particularly dependent on powerful technical tools. In the history of astronomy, scientific progress can in many cases be directly traced to new and more powerful instrumentation. Thus, there is no doubt that the new astronomical instruments, which presently are under construction or in an advanced planning stage, will have a decisive impact on the research field discussed in this book. Some of the future instruments, which are expected to provide particularly important new information on the high-redshift universe, will be discussed below. The section will begin with a summary of the technical specifications and the predicted performance of the new facilities. This will be followed by an assessment of the potential of the new instruments for scientific progress. Finally an outlook beyond the presently planned future instrumentational and scientific possibilities will be attempted.

In contrast to the other sections of this book, only few references to papers on the new instruments will be given. Although there exist many publications on these instruments, the details of most of these projects have been changing rapidly during the past years, and further changes and modifications have to be expected. Even relatively recent publications are already obsolete. Therefore, only the main technical data, which are needed to assess the potential of the future instruments, will be summarized. All the projects mentioned here have extensive web pages (accessible through the major search engines), where comprehensive and (in most cases) up-to-date information on these future facilities can be found.

12.1.1 New Ground-Based Telescopes

12.1.1.1 Optical and Infrared Telescopes

At present the largest optical-infrared telescopes have apertures between 8.2 m and 10.4 m. Currently more than a dozen such "very large telescopes" are in operation. Much of the recent progress of the field of high-redshift galaxies can be traced directly to the development of these powerful facilities. Most important for the

I. Appenzeller, *High-Redshift Galaxies*, Astronomy and Astrophysics Library,
DOI 10.1007/978-3-540-75824-2_12,

progress in the field of high-redshift galaxies have been results obtained with the two 10-m Keck telescopes at Hawaii (www.keckobservatory.org), the ESO VLT (www.eso.org), the Japanese national telescope Subaru (www.naoj.org), and the two Gemini telescopes (www.gemini.edu). However, many of the yet unsolved questions require the light collecting power and the angular resolution of even larger instruments. In view of the experience gained with the 8–10-m class telescopes, today it appears possible to construct ground-based telescopes with apertures 3–5 times larger at affordable costs. Therefore, several large studies or projects have been initiated to realize such "extremely large telescopes" (ELTs). All ELTs are designed for use with adaptive optical systems. At least in the infrared diffraction-limited observations will be possible with these instruments.

Among the most advanced ELT projects is the "Thirty Meter Telescope" (TMT) which is planned jointly by the California Institute of Technology, the University of California, and the Association of Canadian Universities for Research in Astronomy. According to the present plans the TMT will consist of a three-mirror optical system with a 30-m primary mirror. The primary will consist of 492 hexagonal segments. The secondary will be a fully active mirror. A decision on the site of the TMT and many technical details are still open. According to the present timeline the TMT is expected to become operational before 2020. (For more information see: www.tmt.org.)

A second US-lead ELT project is the Giant Magellan Telescope (GMT) (www.gmto.org). The GMT is planned jointly by five US universities, the Carnegie Institution of Washington, and the Australian National University. It will consist of seven 8.4-m mirrors on a common mount and with a common focal plane. Six of these mirrors will surround the central seventh mirror in a "floral" pattern (Fig. 12.1).

Fig. 12.1 Artist's view of the planned Giant Magellan Telescope (GMT). (Credit: Giant Magellan Telescope—Carnegie Observatories)

The GMT will be located at the Las Campanas Observatory in Northern Chile. The manufacturing of the first 8.4-m mirror for the GMT is in progress at the University of Arizona Mirror Lab. According to the present schedule the GMT could see first light by about 2017.

Also relatively advanced is the "European Extremely Large Telescope" (E-ELT) project (Fig. 12.2, www.eso.org). For this telescope a Phase B study is being carried out at the European Southern Observatory, with the support of several national organizations in the ESO countries. According to the present plans the E-ELT will have a segmented primary of 42 m diameter. Depending on the outcome of current feasibility studies, the total optical systems will include either three or five mirrors. In either case one mirror will be fully adaptive. This large adaptive mirror will be among the technological challenges of the E-ELT. With a diameter of about 10 arcmin, the field of view of the E-ELT will be modest. However, due to the planned diffraction-limited resolution, this field will contain a large amount of resolved pixels.

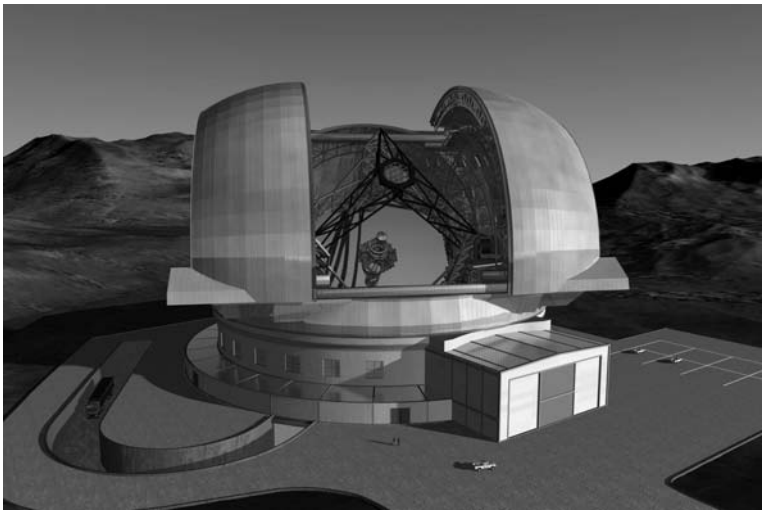

Fig. 12.2 An artist's view of the planned European extremely large telescope. (Courtesy European Southern Observatory)

Some site testing for the E-ELT has been carried out at various locations, but, so far, no decision on the site of the E-ELT has been made. The time line of the project will depend on the funding.

Although not an ELT, the "Large Synoptic Survey Telescope" (LSST) has to be mentioned here, since this future telescope also should be a significant boon to high-redshift galaxy research. Compared to the TMT, the GMT, and the E-ELT the 8.4-m LSST will have a comparatively small aperture. However, its innovative three-mirror optical system will make it possible to do seeing-limited imaging over a 3.5 degree field. Thus, integrated over the field of view, per unit time the LSST will collect more than ten times more photons than the E-ELT. (For details see www.lsst.org.)

An interesting innovative detail of the LSST is the combination of the primary and tertiary optical surfaces in one mirror. The outer part of this mirror will act as the primary, while the inner part (with a different surface shape) is used as tertiary. The site of the LSST will be Cerro Pachón in Northern Chile. First light is envisaged for the year 2014.

To make use of the large FOV of the LSST a matched detector array with about 3×10^9 pixels is being developed. Like the Pan-STARRS telescopes (described in Sect. 11.1), the LSST will be used to continuously map the region of the night-sky which is visible from the telescope site. For this purpose short 10 s exposures will be made at a given sky position. In this way the telescope will image the visible part of the sky about every three nights. The single 10 s images will have a limiting magnitude of about 24. Therefore, these images will contain only the brightest high-redshift galaxies. However, apart from identifying time variable and moving objects (which is the main objective of the LSST), deep sky images will be generated by stacking all images obtained of a certain position. After 3 years this will result in sky maps covering five filter bands with a limiting magnitude of 27. Reaching about 4 magnitudes deeper that the SDSS, these maps will form a highly valuable data base for all fields of astronomy.

Various other very-large or extremely-large telescope projects are under consideration in different countries. However, for these other projects the chances of realization and the technical details are less clear at present. Therefore, their potential for scientific progress cannot yet be assessed.

12.1.1.2 The Atacama Large Millimeter/Submillimeter Array

Among the future new instruments in the millimeter and submillimeter spectral range the Atacama Large Millimeter/Submillimeter Array (ALMA, see www.eso.org) stands out in its potential for progress in the field of high-redshift galaxies. ALMA is an international cooperation between the astronomical communities in Europe, Japan, and the United States. The ALMA array will initially consist of 66 parabolic antennas (Fig. 12.3). Of these antennas, 54 will have a diameter of 12 m. The remaining 12 antennas will have a diameter of 7 m. There are plans to extent the array to (eventually) 64 antenna of 12 m size. The array is under construction on the Chajnantor plain, a plateau about 5000 m above sea level in the Andes of Northern Chile. Because of the high altitude (above a significant fraction of the Earth's atmosphere) and the low water vapor content, Chajnantor is a particularly good site for millimeter and submillimeter astronomy. Several single telescopes, including the 12-m "Atacama Pathfinder Experiment" (APEX) telescope (Fig. 12.4), are already successfully in operation on this plateau. Completion of the ALMA array is envisaged for about 2012.

The 12-m antennas will have reconfigurable baselines ranging from 15 m to 18 km. For changing baselines the 12-m antennas can be moved on specially designed transport vehicles. During the observations the antennas will be parked at pre-selected positions. The digitized signals from the individual antennas will be collected at a correlator site by means of fiber lines. The operating wavelengths of

Fig. 12.3 Two of the prototype antennas of the ALMA array during tests at the VLA site in New Mexico. (Courtesy European Southern Observatory)

Fig. 12.4 The APEX 12-m telescope on the Chajnantor plateau, where the ALMA array will be located. (Courtesy European Southern Observatory)

the array will be between 0.3 and 9.6 mm. At the highest frequency the array will reach an angular resolution of 0.005 arcsec. The flux limit of the instrument (for 60 s integration time and continuum observations) is expected to range between 0.05 mJy at the lowest frequency and about 1 mJy at the highest frequency.

12.1.1.3 New Radio Telescopes

An ambitious future project at radio wavelength is the "Square Kilometer Array" (SKA). It will be an array telescope with a collecting area of 10^6 m^2 operating in the wavelength range 3 cm to 40 m. According to present plans the SKA will consist of a relatively compact inner core and chains of outer stations extending to distances up to 3000 km. The SKA is a truly international project in which (at present) institutions from 19 countries are working together in defining the technical details of this powerful future instrument. The telescope will be located either in South Africa or in Australia.

One of the important new opportunities of the SKA will be the possibility to study the 21-cm line of neutral hydrogen emitted or absorbed by high-redshift objects and by the distant neutral intergalactic medium. As pointed out in Sect. 9.1.1 between a redshifts of $z \approx 200$ and the epoch of reionization the neutral hydrogen gas should be observable by its absorption of the CMB. This effect should be well within reach of the SKA, if the high background at low radio frequencies due to man-made and atmospheric low-frequency radio noise can be adequately suppressed.

Smaller (but still large) low-frequency radio arrays are the "Murchison Wide Field Array" (MWA) in Australia and the "Low Frequency Array" (LOFAR) in Europe. These two arrays, having about 10% of the collecting area of the planned SKA, are close to completion.

The MWA is an international project involving universities and institutes in the US, Australia, and India. It will consist of 8000 dipole antennas. Most of the antennas will be in a 1.5 km core region, with some antennas extending to 3 km. Among the science goals of the MWA is a derivation of the power spectrum of the 21-cm emission near the epoch of cosmic reionization.

LOFAR is a Dutch project with German and other European participation. Its aim is to carry out radio interferometry at wavelengths between about 1.4 and 30 m. The array will consist of a large number of low-cost radio antennas distributed over an area with an diameter of about 300 km, centered on the Netherlands. Its angular resolution will (depending on the wavelength) range between about 1 and 20 arcsec. The technological challenge of LOFAR will be the suppression of interferences by man-made low-frequency radio noise in the most densely populated region of Europe. Among the objectives of LOFAR are again studies of the 21-cm line of high-redshift objects and of the early IGM (e.g., [433]). While such observations may be feasible with LOFAR, estimates show that a large amount of observing time will have to be invested to reach the expected flux levels.

12.1.2 New Space Instrumentation

12.1.2.1 Upgrades to the HST

From the earlier chapters of this book it is clear that the Hubble Space Telescope has played an important role in the rapid progress of the field of high-redshift galaxies.

The HST was launched in 1990. Sometime in the next decade it will be replaced by the JWST (see next subsection). Nevertheless, for several more years the HST will remain the largest space telescope for the UV, visual, and NIR wavelength range. It is expected to provide very valuable data on high redshift galaxies, particularly if the ambitious HST servicing mission No. 4 (at present scheduled for 2009) can be carried out completely and successfully.

Apart from the planned repair of the imaging spectrograph STIS and the Advanced Camera for Surveys (ACS), during the servicing mission two interesting new instruments will be installed at the HST. One of these instruments is the "Wide Field Camera 3" for high-resolution imaging over a wide spectral range, particularly in the UV and near IR. The second new instrument is a new, more sensitive spectrograph ("COS") which is optimized for UV spectroscopy. Both instruments have significant potential for the field discussed here.

12.1.2.2 The James Web Space Telescope

Among all new instrumentation projects the James Web Space Telescope (JWST; Fig. 12.5) is regarded as the most promising tool for making significant progress in the study of high-redshift galaxies. The JWST is a 6.5-m segmented infrared telescope, which will be operated in the Lagrange point L2 of the Earth–Sun system, about 1.5 million km from Earth. The facility is a joint project of NASA, ESA, and the Canadian Space Agency. It is named after a former NASA administrator. Launch is planned for 2013, and a mission duration of 5–10 years is foreseen. Technical details can be found at www.jwst.nasa.gov.

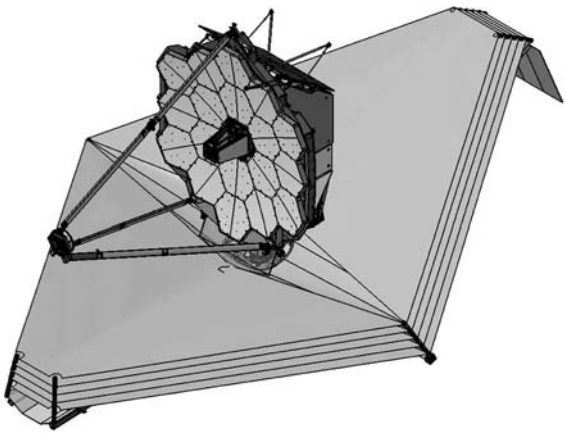

Fig. 12.5 CAD drawing of the James Web Space Telescope. The $\approx$ 6.5-m primary mirror consists of 18 hexagonal segments. The four focal instruments are mounted on the back of the mirror cell assembly. To protect the telescope from sunlight, it is mounted on a 22 m $\times$ 12 m multi-layer sun shield. (Courtesy National Aeronautics and Space Administration)

The wavelength coverage of the JWST will be 0.6–28 μm. To reduce the thermal background the telescope temperature will be kept below 50 K by means of passive radiation cooling. The optical angular resolution will be about 0.1 arcsec. There will be four focal-plane instruments: a near-infrared camera (NIRCam), a near-infrared spectrograph (NIRSpec), a mid-infrared instrument (MIRI), and the Fine Guiding Sensor (FGS) camera system.

NIRCam will be the primary imager for the wavelength range $0.6 < \lambda < 5\,\mu$m. It will consist of two broad-band and intermediate-band imaging modules, each with a 2.16 × 2.16 arcmin field of view. Each module will have a long and a short-wavelength channel. As the light to the two channels is separated by a dichroic, the two channels can be used simultaneously. All detectors are HgCdTe arrays. The pixel scale will be 0.0317 arcsec/pixel in the short-wavelength channels, and 0.0648 arcsec/pix in the long-wavelength channels. The expected point-source sensitivity is about 3.5 nJ for an integration time of 10^5 s and a S/N of 10.

NIRSpec will be the largest instrument on JWST. According to the present design NIRSpec will include multi-object spectroscopy over a 3.4 × 3.4 arcmin field with spectral resolutions of about $R \approx 100$ and $R \approx 1000$. Moreover integral-field and long-slit spectroscopy is being planned with a resolution of $R \approx 3000$. The $R \approx 100$ spectra will cover the wavelength range $0.6\,\mu\text{m} < \lambda < 5\,\mu$m. For the other modes gratings covering the range $1\,\mu\text{m} < \lambda < 5\,\mu$m will be available. Again, HgCdTe detectors will be used. The pixel scale will be about 0.1 arcsec/pixel.

MIRI will provide imaging and spectroscopy at wavelengths between 5 and 27 μm. It will consist of two modules, an imager and a medium resolution spectrograph. Both modules will use 1024 × 1024 pixel Si:As detectors, which will be operated at about 7 K. The low temperature will be reached using solid hydrogen as a coolant.

The imager is designed for broad-band and narrow-band imaging, and (using a prism) low-resolution ($R \approx 100$) spectroscopy in the wavelength range 5 to 10 μm. The pixel scale will be 0.11 arcsec/pixel, the field about 1.3 × 1.7 arcmin.

The medium resolution spectrograph of MIRI will cover the wavelength range of 5–27 μm with a spectral resolution of about $R = 3000$. The instrument will observe simultaneously four small fields of 3.7 × 3.7 arcsec to 7.7 × 7.7 arcsec size, using four integral field units. The pixel scale will range between 0.2 and 0.65 arcsec/pixel.

The main function of the **FGS** is the target acquisition and guiding of the JWST. For this purpose the FGS is equipped with a sensitive and fast NIR camera with a 2.4 × 2.4 arcmin field. In addition to this technical camera the FGS will also include a "tunable filter imager" (TFI) with a 2.2 × 2.2 arcmin field. The TFI will contribute to the scientific capabilities of the JWST by providing monochromatic images with a spectral resolution of $R \approx 100$ in the wavelength range of 1.5–5 μm.

12.1.2.3 The Herschel Space Observatory

Another future space telescope of considerable interest for high-redshift research is the Herschel Space Observatory (usually simply referred to as "HERSCHEL"). HERSCHEL is an infrared satellite developed by the European Space Agency (with

some US participation) for deep observations in the wavelength range $60 < \lambda < 670$ μm. It consists of a 3.5-m lightweight telescope (Fig. 12.6) and the three focal plane instruments HIFI, PACS, and SPIRE.

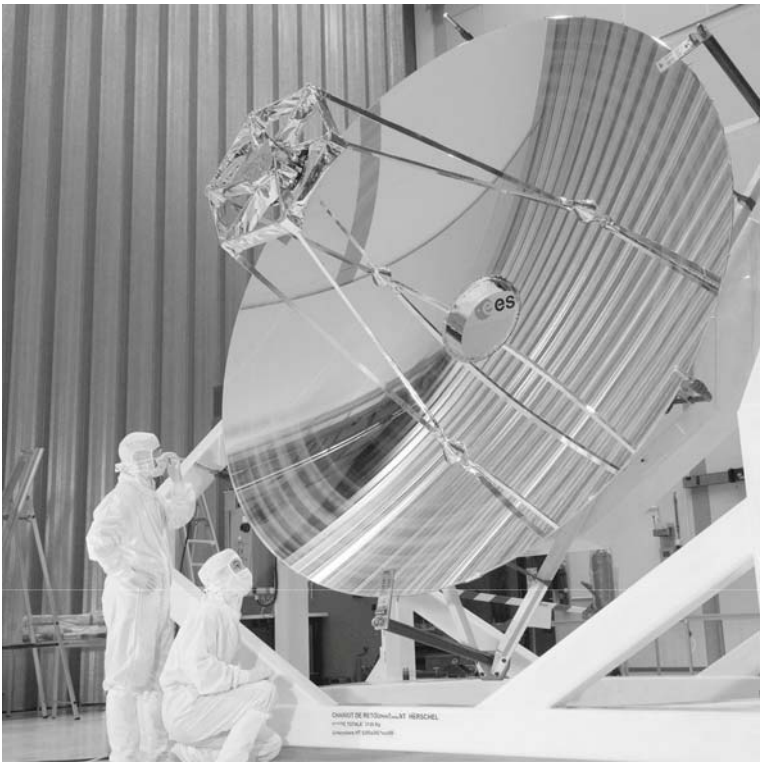

Fig. 12.6 The 3.5-m telescope of the HERSCHEL far-infrared space observatory during laboratory inspection at ESA. (Photo European Space Agency)

HIFI (= Heterodyne Instrument for the Far Infrared) is a very high resolution (up to $R = 5 \times 10^6$) spectrometer for observations of relatively bright objects in the wavelength range of 160–625 μm. As detectors HIFI uses heterodyne receivers with SIS or (at the highest frequencies) hot-electron bolometer (HEB) mixers.

PACS (= Photodetector Array Camera and Spectrometer) will provide images with a 1.75 × 3.5 arcmin FOV at wavelengths between 60 and 210 μm. In spectroscopic mode PACS will image a 50 × 50 arcsec field, resolved into 5 × 5 pixels, with a spectral resolution between 1500 and 2000. The detectors will be bolometer arrays and Ge-Ga photoconductor arrays.

SPIRE (= Spectral and Photometric Imaging Receiver) consists of a an imaging photometer and an imaging Fourier transform spectrometer. All detectors will be bolometers.

The SPIRE photometer has three wavelength bands, centered on 250, 360, and 520 μm. The spectral resolution will be $R \approx 3$ and the FOV about 4×4 arcmin (4×8 arcmin in scanning mode).

The spectrometer has a circular FOV of 2.6 arcmin diameter and spectral resolutions R between 20 and 1000. It will cover the wavelength range between 200 and 670 μm.

The launch of HERSCHEL is foreseen for 2009. HERSCHEL will be launched together with the PLANCK satellite. PLANCK is expected to provide new data on the CMB with an accuracy significantly higher than was reached by WMAP. The PLANCK data will help to constrain more precisely the initial conditions for the formation of high-redshift galaxies.

12.2 New Observational Possibilities and Their Expected Yield

12.2.1 Sensitivity Gains

All the future telescopes, which are under development or under construction, will have significantly larger collecting areas than the corresponding present-day instruments. Thus, per unit time these instruments will collect many more photons. The gain in the photon flux will be $\propto D^2$, where D is the diameter of the telescope aperture. This means for the ELTs discussed above, gain factors between about 10 and 20 relative to the present very large telescopes. According to Eq. (4.10) for extended faint objects, where the sky background dominates the noise (which is normally the case for high-redshift galaxies), the exposure times will become shorter by about these factors. Moreover at a given exposure time and S/N, the flux limit of the observations will be extended $\propto D^{-1}$. For diffraction limited observations of faint point sources, the gain will be higher, since the diffraction pattern of the larger telescopes will be smaller $\propto D^{-1}$, which reduces the sky background contribution $\propto D^{-2}$. Therefore, in this case the exposure times will be reduced $\propto D^{-4}$, and the limiting flux will decrease $\propto D^{-2}$.

At the diffraction-limited resolution of the future ELTs, the high-redshift galaxies will normally be resolved. Thus, the gain for these objects will be closer to the extended-object case. However, high-redshift starburst galaxies are known to have a patchy structure with prominent bright spots. These spots will behave like point sources. Therefore, the actual gains for high-redshift galaxies are expected to range between the extended-object case and the point sources. Expressed in magnitudes the ELTs are expected to reach about 2 to 3 magnitudes deeper. The present limit of optical spectroscopy of about $m_I = 26$ should be extended to at least $m_I = 28$.

The higher light collecting power is also an important advantage of the new space-based facilities for the IR and submm range. At NIR wavelength the gain between the NICMOS camera of the HST and the JWST will be comparable to that between the present ground-based large telescopes and the ELTs. At longer IR wavelengths ($3\,\mu\text{m} \leq \lambda \leq 28\,\mu\text{m}$) the JWST will collect about 60 times more

photons than is possible with the current 85-cm telescope of the Spitzer satellite. In the wavelength overlap region 60–180 μm the HERSCHEL telescope will collect about 17 times more photons than Spitzer. However, since the HERSCHEL telescope will be operated at a higher temperature, the actual sensitivity gain relative to Spitzer will depend on the wavelength and the specific observations.

The gain in sensitivity will make it possible to obtain much more complete samples of high-redshift galaxies. Particularly valuable will be samples of red (dust-obscured and passively evolving) high-redshift galaxies. Observing such galaxies at the highest redshifts where they are expected to occur and at lower flux levels than is possible at present will greatly help clarifying the open questions concerning the early evolution of galaxies.

12.2.2 Extending the Redshift Range

Another important aspect of the new ground-based and space-based new facilities will be extension of the wavelength range in which sensitive observations will be possible. Together with the higher sensitivity discussed above, the extended wavelength range will result in an expansion of the accessible redshift range. NIRCam and NIRSpec on the JWST will, in principle, be capable of observing the Lyα and FUV emission of galaxies to a redshift of about $z = 40$. This will guarantee that the epoch of the formation of the very first galaxies and stars will be in reach of the JWT. While the two NIR instruments hopefully will see the rest-frame UV flux of the first generation of starburst galaxies, MIRI on the JWST will allow us to observe the first evolved stellar populations, as soon as they are formed. The HERSCHEL satellite and the ALMA array will make it possible to study the molecular and dust content of larger and more representative samples of high-redshift galaxies than has been possible so far. However, if the predictions are correct, with the currently expected sensitivities, detailed studies of the dynamics of high-redshift galaxies using molecular lines will still be difficult even with ALMA [211].

With HERSCHEL and ALMA an almost continuous photometric coverage of the wavelength range between 60 μm and 7 mm will be possible. This will be important for deriving dust temperatures, the dust composition, and the dust masses of the high-redshift galaxies.

12.2.3 Gains from the Improved Angular Resolution

A further advantage of the new ground-based and space-based facilities will be their improved angular resolution. Although high-redshift galaxies are normally resolved even with present-day telescopes, their sizes are often close to the resolution limit. This makes reliable studies of their morphology and their geometric and kinematic properties difficult. With the future, larger space-based telescopes and adaptive-optics observations on ground-based ELTs the resolution will increase

$\propto D$. In the submillimeter and millimeter wavelength range ALMA will increase the resolution compared to existing arrays by factors of the order of 10. Moreover, the good UV-plane coverage and large flux collecting area of ALMA will result in much better image information than can be obtained with existing arrays at the ALMA wavelengths. Among the important new observational results expected from ALMA are spatially resolved CO and continuum observations of bright dusty high-redshift starburst galaxies.

12.3 The Long-Term Future

In view of past experience it seems likely that the large new ground-based instruments, which are under construction or in an advanced planning stage, will be operated for many decades, before they are replaced by technologically superior facilities. Technical upgrades will almost certainly increase the performance of these instruments significantly over the presently foreseen specifications. And new detectors and focal plane instruments will make the new large ground-based telescopes more efficient than planned initially.

On the other hand, all the new large space-based facilities will be operated near the Lagrange point L2, where a refurbishment and upgrades are impossible, or at least difficult and expensive. Therefore, when severe problems occur, or when their components become obsolete, these instruments will have to be replaced. Their useful lifetime will depend on their technical health, the lifetime of the focal instruments, and the development of the important scientific questions. In view of the high development costs of such facilities, they may well be replaced once or more times by similar or slightly modified instruments. Eventually the facilities in L2 may be replaced by observatories on the Moon, if a permanent Moon base were to be established. Although expensive to operate, a site on the surface of the moon could combine some advantages of present space-based and ground-based facilities. Moreover, radio observations from the Moon would eliminate the attenuation of radio waves by the terrestrial ionosphere, and at long radio wavelengths the far side of the Moon would provide optimal shielding from terrestrial radio noise.

12.3.1 Medium and Long-Term Projects in Space

Among the long-term initiatives under consideration at NASA and ESA are several missions of significant interest for the field of high-redshift galaxies. Among the most promising proposals for future projects are the XEUS (standing for "X-ray Evolving Universe Spectroscopy") project being discussed at ESA and the Constellation-X proposal under consideration at NASA. According to present plans XEUS will have a collecting area of about $5\,m^2$ for 1 kV photons and a resolution of a few arcsec. Its high sensitivity would be of great value for studying the X-ray emission of high-redshift starburst galaxies and distant AGN.

XEUS and Constellation-X both aim at an significant improvement of the sensitivity of future X-ray observation. There are discussions of possibly joining the two projects into one powerful future X-ray facility.

Another project being considered by ESA is the EUCLID optical survey mission. Its main objective is a precise all-sky map of the galaxy distribution at redshifts ≤ 2. This map will be used to study weak lensing and the baryonic oscillations up to this redshift. Although the universe at $z > 2$ is outside the original scope of EUCLID, this mission is expected to yield also interesting data on the high-redshift universe.

Important new observational results could also be expected from the proposed SPICA infrared satellite. This proposal, which is being studied at ESA, aims at photometry and spectroscopy in the wavelength range from 5 to 210 μm. It would close the gap in the wavelength coverage between the JWST and HERSCHEL.

A future astronomical space mission which could have a particularly strong impact on the field of high-redshift galaxies is the "Spectroscopic All-Sky Cosmic Explorer" (SPACE) project [100]. This mission has been proposed as a concept study for the 2015–2025 planning cycle of ESA. The present plans foresee a 1.5-m space-based telescope optimized for IR spectroscopy. The aim of SPACE is to obtain spectra of $\approx 5 \times 10^8$ galaxies with redshifts ≤ 2, and of several 10^6 galaxies with redshifts $2 < z < 10$.

12.3.2 Advances in the Detector Technologies

During part of the second half of the 20th century advances in the detector technology were more important for the progress in astronomy than new telescopes. This situation may well occur again in the future. As pointed out in Sect. 4.3 imaging detectors used for X-ray astronomy provide imaging as well as spectral information by directly measuring the energy of the recorded photons. Using superconducting junction arrays, this is also possible at optical wavelength (e.g., [250]). Technical difficulties of producing large arrays of such junctions have, so far, limited this technique to a few special applications. If future technical developments could make it possible to produce large-format superconducting imaging detectors, astronomy would greatly profit from such a development. By replacing filter photometry, it would result in a large efficiency gain for many studies of the high-redshift objects.

Technical progress in the last decade of the 20th century opened up the research field of high-redshift galaxies. The large new instruments, which are under construction, will make it possible to clarify many of the still open questions. But there will remain problems which are out of reach of all the present projects and ideas. Very likely these problems will eventually be solved using powerful technical tools which are well beyond our present imagination.

References

1. Abel, T., Bryan, G., Norman, M. L. 2000, ApJ **540**, 39 113
2. Abel, T., Bryan, G., Norman, M. L. 2002, Science **295**, 93 113
3. Adelberger, K. L., Steidel, C. C., Pettini, M., et al. 2005, ApJ **619**, 697 272
4. Adelberger, K. L., Steidel, C. C., Shapley, A. E., et al. 2003, ApJ **584**, 45 238, 287
5. Adelberger, K. L., Steidel, C. C., Shapley, A. E., et al. 2004, ApJ **607**, 226 156, 157
6. Ahn, S.-H. 2004, ApJ **601**, L25 254
7. Ahn, S.-H., Lee, H.-W., Lee, H. M. 2001, ApJ **554**, 604 254
8. Ahn, S.-H., Lee, H.-W., & Lee, H. M., 2003, MNRAS **340**, 863 254
9. Ajiki, M., Mobasher, B., Taniguchi, Y., et al. 2006, ApJ **638**, 596 248
10. Ajiki, M., Taniguchi, Y., Fujita, S. S., et al. 2003, AJ **126**, 2091 248
11. Ajiki, M., Taniguchi, Y., Fujita, S. S., et al. 2004, PASJ **56**, 597 248
12. Akiyama, M., Minowa, Y, Kobayashi, N., et al. 2008 ApJS **175**, 1 211, 214
13. Alexander, D. M., Bauer, F. E., Brandt, . N., et al. 2003, AJ **126**, 539 170
14. Appenzeller, I. 1970, A&A, **5**, 355 20
15. Appenzeller, I. 1970, A&A, **9**, 216 20
16. Appenzeller, I., Bender, R., Böhm, A., et al. 2000, The Messenger **100**, 44 168
17. Appenzeller, I., Bender, R., Böhm, A., et al. 2004, The Messenger **116**, 18 210
18. Appenzeller, I., Mandel, H., Krautter, J., et al. 1995, ApJ **439**, L33 68
19. Appenzeller, I., Stahl, O., Tapken, C., et al. 2005, A&A **435**, 465 61
20. Appenzeller, I., Tscharnuter, W. 1974, A&A **30**, 423 20
21. Aretxaga, I., Hughes, D. H., Coppin, Kristen, et al. 2007, MNRAS **379**, 1571 209
22. Astier, P., Guy, J., Regnault, N., et al. 2006, A&A **447**, 31 116
23. Audley, M. D., Holland, D., Atkinson, M., et al. SCUBA-2: A Large-Format CCD-Style Imager for Submillimeter Astronomy. In: A. P. Lobanov, J. A. Zensus, C. Cesarsky, P. J. Diamond (eds) *Exploring the Cosmic Frontier* (Springer, Berlin Heidelberg, New York 2007) p. 45 133, 135
24. Baade, W. 1944, ApJ **100**, 137 42
25. Baars, J. W. M. *The Paraboloidal Reflector Antenna in Radio Astronomy and Communication* (Springer Science+Business Media, LLC, New York 2007) 123
26. Bahcall, J. N., Kirhakos, S., Saxe, D. H., Schneider, D. P. 1997, ApJ **479**, 642 263
27. Baker, A. J., Lutz, D., Genzel, R., et al. 2001, A&A **372**, L37 242
28. Baker, A. J., Tacconi, L. J., Genzel, R., et al. 2004, ApJ **604**, 125 197, 235
29. Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, ApJ **600**, 681 40, 41
30. Bardeen, J. M., Bond, J. R., Kaiser, N., Szalay, A. S. 1986, ApJ **304**, 15 93
31. Barkana, R., Loeb, A. 1999, ApJ **523**, 54 287
32. Baum, W. A. 1962, Proc. IAU Symposium **15**, 390 178
33. Becker, R. H., Fan, X., White, R. L., et al. 2001, AJ **122**, 2850 280
34. Becker, R. H., White, R. L., Helfand, D. J., 1995, ApJ **450**, 559 165
35. Beckwith, S. V. W, Stiavelli, M., Koekemoer, A. M., et al., 2006, AJ **132**, 1729 139, 166, 212
36. Beers, C. B., Christlieb, N. 2005, Ann. Rev. Astron. Astrophys. **43**, 531 115, 301

37. Begelman, M. C., Volonteri, M., Rees, M.J. 2006, MNRAS **370**, 289 303
38. Beletic, J. E., Beletic, J. W., Amico, P.(eds.) *Scientific Detectors for Astronomy* (Springer, Berlin Heidelberg New York 2006) 134
39. Bell, E. F., Wolf, C., Meisenheimer, K., et al. 2004, ApJ **608**, 752 46, 178
40. Bell, E. F., Naab, T., McIntosh, D. H., et al. 2006, ApJ **640**, 241 305
41. Bell, E. F., Phleps, S., Somerville, R. S., et al. 2006, ApJ **652**, 270 305
42. Bely, P. Y. (ed.) *The Design and Construction of Large Optical Telescopes* (Springer, Berlin Heidelberg, New York 2003) 124
43. Bender, R. 1988, A&A **193**, L7 33
44. Bender, R., Döbereiner, S., Möllenhoff, C. 1988, A&A Suppl. Ser. **74**, 385 35
45. Bender, R., Saglia, R. P. Elliptical Galaxies: Detailed Structure, Scaling Relations and Formation. In: D. Merritt et al. (eds) *Galaxy Dynamics*, ASP Conference Series Vol. 182. (ASP 1999) p. 113 33, 62
46. Benjamin, J., Heymans, C., Semboloni, E., et al. 2007, MNRAS **381**, 702 107
47. Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003, ApJS **148**, 1 82, 102, 103
48. Berger, E., Penprasse, B. E., Cenko, S. B., et al. 2006, ApJ **642**, 979 267
49. Bertin, E. Arnouts, S. 1996, A&AS, **117**, 393 138
50. Bertoldi, F., Carilli, C. L., Cox, P., et al. 2003, A&A **406**, L55 264
51. Bessell, M. S. 1990, PASP **102**, 1181 141
52. Bessell, M. S. 2005 Ann. Rev. Astron. Astrophys. **43**, 293 141, 142
53. Blandford, R. D. Physical Processes in Active Galactic Nuclei. In: T. J. L. Courvoisier and M. Mayor (eds) *Active Galactic Nuclei* (Springer, Berlin Heidelberg, New York 1990) p. 161 30
54. Böhm, A., Ziegler, B. L. 2007, ApJ **668**, 846 176
55. Bolton, J. S., Haehnelt, M. G. 2007, MNRAS **381**, 35 286
56. Bolton, J. S., Haehnelt, M. G. 2007, MNRAS **382**, 325 286
57. Bolzonella, M., Miralles, J.-M., Pello, R. 2000, A&A **363**, 476 178
58. Borch, A., Meisenheimer, K., Bell, E. F., et al. 2006, A&A **453**, 869 187
59. Born, M., Wolf, E. *Principles of Optics* 6th edn (Pergamon Press , Oxford 1980) 130
60. Börner, G. *The Early Universe*, 4th edn (Springer, Berlin Heidelberg, New York 2003) 41, 81, 108
61. Bouché, N., Cresci, G., Davies, R., et al. 2007, ApJ **671**, 303 240, 241
62. Bouwens, R. J., Illingworth, G. D. 2006, Nature **443**, 189 208
63. Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., Franx, M. 2004, ApJ **611**, L1 214
64. Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., Franx, M. 2006, ApJ **653**, 53 167, 183, 208, 212, 214
65. Bouwens, R. J., Illingworth, G. D., Franx, M., Ford, H. 2007, ApJ **670**, 928 214, 215, 217, 219, 286
66. Bouwens, R. J., Illingworth, G. D., Franx, M., Ford, H. 2008, ApJ **686**, 230 216
67. Boyce, P. J., Disney, M. J., Blades, J. C., et al. 1998, MNRAS **298**, 121 263
68. Bradley, L. D., Bouwens, R. J., Illingworth, G. D., et al. 2008, ApJ **678**, 647 10, 173, 299
69. Brandenberger, R., Kahn, R. 1984, Phys. Lett. **141B**, 317 87
70. Brandt, W. N., Hasinger, G. 2005, Ann. Rev. Astron. Astrophys. **43**, 827 170, 282
71. Brinchmann, J., Pettini, M., Charlot, S. 2008, MNRAS **385**, 769 300
72. Bromm, V., Coppi, P. S., Larson, R. B. 1999, ApJ **527**, L5 113
73. Bromm, V., Coppi, P. S., Larson, R. B. 2002, ApJ **564**, 23 113
74. Bromm, V., Larson, R. B. 2004, Ann. Rev. Astron. Astrophys. **42**, 79 113, 114
75. Bromm, V., Loeb, A. 2003, Nature **425**, 812 115
76. Brooks, A. M., Governato, F., Booth, C. M., et al. 2007, ApJ **655**, L17 297
77. Brunner, R. J., Hall, P. B., Djorgovski, S. G., et al. 2003, AJ **126**, 53 60
78. Bruzual, G., Charlot, S. 2003, MNRAS **344**, 1000 183, 260
79. Bunker, A. J., Stanway, E. R., Ellis, R. S., McMahon, R. G. 2004, MNRAS **355**, 374 219
80. Burles, S., Nollett, K. M., Turner, M. S. 2001, ApJ **552**, L1 82
81. Burstein, D., Faber, S. M., Gaskell, C. M., Krumm, N. 1984, ApJ **287**, 586 196
82. Burstein, D., Heiles, C. 1982, AJ **87**, 1165 150
83. Caldwell, J. A. R., McIntosh, D. H., Rix, H.-W., et al. 2008, ApJS **174**, 136 166
84. Calzetti, D., Kinney, A. L., Storchi-Bergmann, T. 1994, ApJ **429**, 582 181
85. Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ **533**, 682 181

86. Camenzind, M. *Compact Objects in Astrophysics* (Springer, Berlin Heidelberg, New York 2007) 17, 24, 29
87. Capak, P., Carilli, C. L., Lee, N., et al. 2008, ApJ **681**, 53 258, 304
88. Carilli, C. L., Murayama, T., Wang, R., et al. 2007, ApJS **172**, 518 250
89. Carilli, C. L., Neri, R., Wang, R., et al. 2007, ApJ **666**, L9 264
90. Carilli, C. L., Walter, F., Bertoldi, F., et al. 2004, AJ **128**, 997 264
91. Carswell, R. F., Morton, D. C., Smith, M. G., et al. 1984, ApJ **278**, 486 202
92. Chabrier, G. 2001, ApJ **554**, 1274 13
93. Chakrabarti, S. K. (ed.) *Observational Evidence for Black Holes in the Universe* (Kluwer Academic Publishers, Dordrecht 1999) 30
94. Chapman, S. C., Blain, A. W., Smail, I., et al. 2005, ApJ **622**, 772 209
95. Charles, P.: Black Holes in our Galaxy. In: S. K. Chakrabarti, (ed.) *Observational Evidence for Black Holes in the Universe* (Kluwer Academic Publishers, Dordrecht 1999) 24
96. Charlot, S., & Fall, S. M. 1993, ApJ **415**, 580 249
97. Chary, R.-R., Teplitz, H.I., Dickinson, M. E., et al. 2007, ApJ **665**, 257 276, 295
98. Chernyakova, M., Neronov, A., Courvoisier, T. J.-L., et al. 2007, A&A **465**, 147 57
99. Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, A&A **482**, 21 260, 263, 304
100. Cimatti, A., Robberto, M., Baugh, C., et al. 2009, Experimental Astronomy, **23**, 39 331
101. Coc, A., Vangioni-Flam, E., Descouvemont, P. 2004, ApJ **600**, 544 82
102. Coelho, P., Bruzual, G., Charlot, S., et al. 2007, MNRAS **382**, 498 183
103. Cohen, J. G., Cowie, L. L., Hogg, D. W., Songaila, A. 1996, ApJ **471**, L5 269
104. Cole, S., Norberg, P., Baugh, C. M., et al. 2001, MNRAS **326**, 255 222
105. Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS **362**, 505 104
106. Colless, M., et al. (The 2dFGRS Team) 2001, MNRAS **328**, 1039 63, 148
107. Combes, F., Boissé, P., Mazure, A., Blanchard, A. *Galaxies and Cosmology*, 2nd edn (Springer, Berlin Heidelberg, New York 2002) 32
108. Condon, J. J., Cotton, W. D., Greisen, E., W., et al. 1998, AJ **115**, 1693 165
109. Connolly, A., Scranton, R., Johnston, D., et al. 2002, ApJ **579**, 42 271
110. Conroy, C. Shapley, A. E., Tinker, J. L., et al. 2008, ApJ **679**, 1192 305
111. Conselice, C. J., Bundy, K, Trujillo, I., et al. 2007, MNRAS **381**, 962 262
112. Coppin, K. E. K., Swinbank, A. M., Neri, R., et al. 2007, ApJ **665**, 936 197, 235
113. Courvoisier, T. J.-L. 1998, Astronomy and Astrophysics Reviews **9**, 1 57
114. Cowie, L. L., Hu, E. M. 1998, AJ **115**, 1319 248
115. Crowther, P. A. 2007, Ann. Rev. Astron. Astrophys. **45**, 177 14, 17
116. Crowther, P. A., Prinja, R. K., Petini, M., Steidel, C. C. 2006, MNRAS **368**, 895 194
117. Curtis, H.D. 1917, PASP **29**, 206 7
118. Daddi, E., Cimatti, A., Renzini, A, et al. 2004, ApJ **617**, 746 156, 158, 259
119. Dannerbauer, H., Daddi, E., Onodera, M. 2006, ApJ **637**, L5 242, 256
120. Dannerbauer, H., Walter, F., Morrison, G. 2008, ApJ **673**, L127 258
121. Davis, M., et al. 2003, Proc. SPIE **4834**, 161 166
122. Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, ApJ **660**, L1 169
123. Dawson, S., Spinrad, H., Stern, D., et al. 2002, ApJ **570**, 92 240, 254
124. Dawson, S., Rhoads, J. E., Malhotra, S., et al. 2004, ApJ **617**, 707 248, 250
125. Dawson, S., Rhoads, J. E., Malhotra, S., et al. 2007, ApJ **671**, 1227 209
126. de Bruijne, J. H. J., Hoogerwerf, R., de Zeeuw, P. T. 2001, A&A **367**, 111 19, 20
127. Deffayet, C., Dvali, G., Gabadadze, G. 2002, Phys. Rev. D **65**, 4023 71
128. de Sitter, W. 1918, MNRAS **78**, 3 9
129. de Vaucouleurs, 1959, Handbuch der Physik **53**, 275 38
130. Dey, A., Bian, C., Soifer, B., et al. 2005, ApJ **629**, 654 255
131. Dickinson, M., 1998, STScI Symp. Ser. **11**, 219 155
132. Dickinson, M., Giavalisco, M., the GOODS Team: The Great Observatories Origins Deep Survey. In: R.Bender, A. Renzini (eds) *The Mass of Galaxies at Low and High Redshift* (Springer, Berlin Heidelberg, New York 2003) p. 324 169

133. Dickinson, M., Papovich, C., Ferguson, H. C., Budavári, T. 2003, ApJ **587**, 25 222
134. Dicus, D. A., Letaw, J. R., Teplitz, D. C., Teplitz, V. L. 1983, ApJ **252**, 1 116
135. Dicus, D. A., Letaw, J. R., Teplitz, D. C., Teplitz, V. L. 1983, Scientific American **248**, 90 116
136. Dietrich, M., Hamann, F., Apenzeller, I., Vestergaard, M. 2003, ApJ **596**, 817 265
137. Di Matteo, T., Colberg, J., Springel, V., et al. 2008, ApJ **676**, 33 303
138. Djorgovski, S., Davis, M. 1987, ApJ **313**, 59 35
139. Djorgovski, S. G., Pahre, M. A., Bechtold, J., & Elston, R. 1996, Nature **382**, 234 290
140. Djorgovski, S. G., Spinrad, H., McCarthy, P. 1985, ApJ **299**, 1 10, 158, 206
141. Doherty, M., Bunker, A. J., Ellis, R. S., McCarthy, P. J. 2005, MNRAS **361**, 525 259, 260
142. Dopita, M. A., Sutherland, R. S. *Astrophysics of the Diffuse Universe* (Springer, Berlin Heidelberg, New York 2003) 189, 190, 191
143. Dow-Hygelund, C. C., Holden, B. P., Bouwens, G. D., et al. 2005, ApJ **630**, L137 221
144. Dow-Hygelund, C. C., Holden, B. P., Bouwens, G. D., et al. 2007, ApJ **660**, 47 249, 250, 251
145. Dressler, A., Lynden-Bell, D., Burnstein, D., et al. 1987, ApJ **313**, 42 35
146. Drory, N., Alvarez, M. 2008, ApJ **680**, 41 217, 218
147. Drory, N., Bender, R., Feulner, B., et al. 2004, ApJ **608**, 742 187, 222
148. Drory, N., Salvato, M., Gabasch. A. 2005, ApJ **619**, L131 222
149. Dunlop, J. S., Cirasuolo, M., McLure, R. J. 2007, MNRAS **376**, 1054 295
150. Dye, S., Eales, S. A., Aretxaga, I., et al. 2008, MNRAS **386**, 1107 295
151. Efstathiou, G. Cosmological Perturbations. In: J. A. Peacock, A. F. Heavens, and A. T. Davies (eds) *Physics of the Early Universe* (SUSSP Publications, Edinburgh 1990) p. 361 110
152. Efstathiou, G., Bernstein, G., Katz, N. 1991, ApJ **380**, L47 271
153. Efstathiou, G., Ellis, R. S., Peterson, B. A. 1988, MNRAS **232**, 431 183
154. Eisenstein, D. J., Loeb, A. 1995, ApJ **443**, 11 303
155. Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ **633**, 560 106
156. Elsner, F., Feulner, G., Hopp, U. 2008, A&A **477**, 503 220, 222, 295, 296
157. Eminian, C., Kauffmann, G., Charlot, S., et al. 2008, MNRAS **384**, 930 317
158. Epchtein, N., de Batz, B., Copet, E., et al. 1994, Astr. Sp. Science **217**, 9 164
159. Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, ApJ **644**, 813 223, 230, 231
160. Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, ApJ **646**, 107 235
161. Eyles, L. P., Bunker, A. J., Ellis, R. S., et al. 2007, MNRAS **374**, 910 299
162. Faber, S.M., Gallagher, J.S. 1979, Ann. Rev. Astron. Astrophys. **17**, 135 69
163. Faber, S. M., Jackson, R. 1976, ApJ **204**, 668 35
164. Falomo, R., Treves, A., Kotilainen, J. K., et al. 2008, ApJ **673**, 694 264
165. Fan, X., Carilli, C. L., Keating, B. 2006, Ann. Rev. Astron. Astrophys. **44**, 415 285
166. Fan, X., Strauss, M. A., Schneider, D.P., et al. 1999, ApJ **526**, L77 10
167. Fan, X., Strauss, M. A., Schneider, D.P., et al. 2003, ApJ **125**, 1649 10, 281
168. Fan, X., Strauss, M. A., Richards, G. T. et al. 2006, AJ **131**, 1203 10
169. Ferguson, H. C., Dickinson, M., Giavalisco, M., et al. 2004 ApJ **600**, L107 214
170. Ferguson, H. C., Dickinson, M., Williams, R. 2000, Ann. Rev. Astron. Astrophys. **38**, 667 167, 212
171. Finkelstein, S. L., Rhoads, J. E., Malhotra, S., et al. 2007, ApJ **660**, 1023 251
172. Finkelstein, S. L., Rhoads, J. E., Malhotra, S., et al. 2007, AAS **211**, 2405 253
173. Fioc, M., Rocca Volmerange, B. 1997, A&A **326**, 950 183
174. Fontana, A., Donnarumma, I., Vanzella, E., et al. 2003, ApJ **594**, L9 222
175. Fontana, A., Salimbeni, S., Grazian, A., et al. 2006, A&A **459**, 745 222
176. Förster Schreiber, N. M., Genzel, R., Lehnert, M. D., et al. 2006, ApJ **645**, 1062 240
177. Frank, S., Appenzeller, I., Noll, S., Stahl, O. 2003, A&A **407**, 473 201
178. Franx, M., Labbé, I., Rudnick, G., et al. 2003, ApJ **587**, L79 258
179. Fry, J. N. 1996, ApJ **461**, L65 273
180. Freudling, W., Møler, P., Patat, F., et al. 2007, The Messenger **128**, 13 136
181. Friedmann, A., A. 1922, Zeitschr. Physik **10**, 377 9, 76
182. Friedmann, A., A. 1924, Zeitschr. Physik **21**, 326 9, 76
183. Fruchter, A. S., Levan, A. J., Strolger, L., et al. 2006, Nature **441**, 463 267
184. Frye, B. L., Coe, D., Bowen, D. V., et al. 2007, ApJ **665**, 921 171

185. Fryer, C.L., Woosley, S. E., Heger, A. 2001, ApJ **550**, 372 25
186. Fu, L., Semboloni, E., Hoekstra, A., et al. 2008, A&A **479**, 9 107
187. Fukugita, M., Peebles, P. J. E. 2004, ApJ **616**, 643 72
188. Furusawa, H., Kosugi, G., Akiyama, M., et al. 2008, ApJS, **176**, 1 169
189. Gabasch, A., Bender, R., Seitz, S., et al. 2004, A&A **421**, 41 178, 207
190. Gabasch, A., Goranova, Y., Hopp, U., et al. 2008, MNRAS **383**, 1319 136, 216
191. Gaitskell, R. J. 2004, Ann. Rev. Nucl. Part. Sci. **54**, 315 70
192. Gallagher, J. S., Hunter, D. A., Bushouse, H. 1989, AJ **97**, 700 189
193. Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M. 2006, MNRAS **370**, 1106 46
194. Gallerani, S., Ferrara, A., Fan , X., Choudhury, T. R. 2008, MNRAS **386**, 359 281, 285
195. Gawiser, E., Franke, H., Lai, K., et al. 2007, ApJ **671**, 278 209
196. Gawiser, E., van Dokkum, P. G., Gronwall, C., et al. 2006, ApJ **642**, L13 247, 249, 250, 251
197. Gawiser, E., van Dokkum, P. G., Herrera, D., et al. 2006, ApJS **162**, 1 166
198. Geach, J. E., Smail, I., Chapman, S. C., et al. 2007, ApJ **655**, L9 255
199. Geller, M. J., Huchra, J. 1989, Science **246**, 897 66
200. Genzel, R., Cesarsky, C. J. 2000, Ann. Rev. Astron. Astrophys. **38**, 761 52
201. Genzel, R., Tacconi, L.J., Eisenhauer, F., et al. 2006, Nature **442**, 786 240, 241
202. Giacconi, R., Zirm, A., Wang, J. X., et al. 2002, ApJS **139**, 369 170
203. Giavalisco, M. 2002, Ann. Rev. Astron. Astrophys. **40**, 579 155, 176, 206, 218
204. Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ **600**, L93 169
205. Giavalisco, M., Steidel, C. C., , Adelberger, K. L., et al. 1998, ApJ **503**, 543 271
206. Glazebrook, K., Blake, C., Couch, W., et al. 2007, in "Cosmic Frontiers", ASP Conference Series **379**, 72 312
207. Gonzales Delgado, R. M., Cerviño, M., Martins, L. P., et al. 2005, MNRAS **357**, 945 183
208. Grazian, A., Fontana, A., de Santis, C., et al. 2006, A&A **449**, 951 209, 220
209. Greif, T. H., Johnson, J. L., Klessen, R. S., Bromm, V. 2008, MNRAS **387**, 1021 110, 111, 112, 113, 318
210. Greve, T. R., Bertoldi, F., Smail, I., et al. 2005, MNRAS **359**, 1165 197
211. Greve, T. R., Sommer-Larsen, J. 2008, A&A **480**, 335 197, 329
212. Gronwall, C., Ciardullo, R., Hickey, T., et al. 2007, ApJ **667**, 79 248
213. Groth, E. J., Peebles, P. J. E. 1977, ApJ **217**, 385 274
214. Guhathakurta, P., Tyson, J. A., Majewski, S. R. 1990, ApJ **357**, 9 156
215. Gunn, J. E., Peterson B. A. 1965, ApJ **142**, 1633 279
216. Hall, P.B., Anderson, S. F., Strauss, M. A. 2002, ApJS, **141**, 267 60
217. Halliday, C., Daddi, E., Cimatti, A., et al. 2008, A&A **479**, 417 233
218. Hamann, F., Ferland, G. 1999, Ann. Rev. Astron. Astrophys. **37**, 487 265
219. Hammer, F., Flores, H., Schaerer, D., et al. 2006, A&A **454**, 103 267
220. Häring, N., Rix, H.-W. 2004, ApJ, **604**, 89 56
221. Häußler, B., McIntosh, D. H., Barden, M., et al. 2007, ApJS **172**, 615 139
222. Häußler, B., Bell, E. F., Barden, M., et al. 2007, Proc. IAU Symp. **235**, 102 44
223. Haensel, P., Potekhin, A. Y., Yakovlev, D.G. *Neutron Stars 1* (Springer, Berlin Heidelberg, New York 2007) 17, 24
224. Heckman, T. M., Hoopes, C. G., Seibert, M., et al. 2005, ApJ **619**, L35 305
225. Heckman, T. M., Robert, C., Leitherer, C., et al. 1998, ApJ **503**, 646 42, 194, 232, 243
226. Heckman, T. M., Sembach, K. R., Meurer, G. R., et al. 2001, ApJ **558**, 56 238
227. Heger, A., Fryer, C. L., Woosley, S. E., et al. 2003, ApJ **591**, 288 25
228. Heidt, J., Appenzeller, I., Gabasch, A. et al. 2003, A&A **398**, 49 138, 139, 168
229. Hildebrand, R. H. 1983, QJRAS **24**, 267 199
230. Hillebrandt, W., Niemeyer, J. C. 2000, Ann. Rev. Astron. Astrophys. **38**, 191 27
231. Hinshaw, G., Nolta, M. R., Bennett, C. L. et al. 2007, ApJS **170**, 288 100
232. Hoekstra, H., Mellier, Y., van Waerbeke, L. et al. 2006, ApJ **647**, 116 107
233. Holtzman, J. A., Burrows, C. J., Casertano, S. et al. 1995, PASP **107**, 1065 141
234. Horton, A., Parry, I., Bland-Hawthorn, J., et al. 2004, SPIE **5492**, 1022 313
235. Hu, E. M., Cowie, L. L. 1987, ApJ **317**, L7 10
236. Hu, E. M., McMahon, R. G. 1996, Nature **382**, 231 246

237. Hu, E. M., Cowie, L. L.; McMahon, R., G. 1998, ApJ **502**, L99 246, 247
238. Hu, E. M., Cowie, L. L., Capak, P., et al. 2004, AJ, 127, 563 248, 249
239. Hubble, E. P. 1925, Popular Astronomy **33**, 252 7
240. Hubble, E. P. 1925, ApJ **62**, 409 7
241. Hubble, E. P. 1929, Proc. Nat. Acad. Sci. **15**, 168 8, 9
242. Hubble, E. *The Realm of the Nebulae* (Yale University Press, New Haven 1936) 32, 33
243. Humason, M. L., Pease, F., G. 1929, PASP **41**, 246 10
244. Hunt, L. K., Maiolino, R. 2005, ApJ **626**, L15 179, 256
245. Hunt, M. P., Steidel, C. C., Adelberger, K. L., Shapley, A. E. 2004, ApJ, **605**, 625 285
246. Ivison, R. J., Smail, I., Dunlop, J. S., et al. 2005, MNRAS **364**, 1025 209, 258
247. Iwata, I., Inoue, A. K., Matsuda, Y., et al. 2009, ApJ **692**, 1287 284
248. Iye, M., Ota, K., Nobunari, K., et al. 2006, Nature **433**, 186 209
249. Jahnke, K., Sánchez, S. F., Wisotzki, L., et al. 2004, ApJ **614**, 568 264
250. Jakobsen, P. 1999, in: Ultraviolet-Optical Space Astronomy Beyond HST, ASP Conference Series **164**, 397 331
251. Jappsen, A.-K., Glover, S. C. O., Klessen, R. S., Mac Low, M.-M. 2007, ApJ **660**, 1332 115
252. Jiang, L., Fan, X., Annis, J., et al. 2008, AJ, **135**, 1057 285, 286
253. Johnson, H. L. 1966, Ann. Rev. Astron. Astrophys. **4**, 193 140, 141
254. Johnson, J. L., Greif, T. H., Bromm, V. 2008, Proc. IAU Symposium **250**, 471 113, 294, 318
255. Johnson, J. L., Greif, T. H., Bromm, V. 2008, MNRAS **388**, 26 318
256. Jorgensen, I., Franx, M., Kjaergaard, P. 1996, MNRAS **280**, 167 35
257. Karlsson, T., Johnson, J. L., Bromm, V. 2008, ApJ **679**, 6 301
258. Kashikawa, N., Shimasaku, K., Matthew, A., et al. 2006, ApJ **648**, 7 209
259. Kashikawa, N., Shimasaku, K., Yasuda, N., et al. 2004, Publ. Astr. Soc. Japan **56**, 1011 168
260. Kashlinsky, A., Arendt, R., Mather, J., Moseley, S. 2005, Nature **438**, 45 300
261. Kashlinsky, A., Arendt, R., Mather, J., Moseley, S. 2007, ApJ **666**, L1 300
262. Kauffmann, G. 1996, MNRAS **281**, 475 294
263. Kauffmann, G., White, S. D. M., Guderdoni, G. 1993, MNRAS **264**, 201 97
264. Kawaler, S. D., Novikov, I., Srinivasan, G. *Stellar Remnants* (Saas-Fee Advanced Course 25) (Springer, Berlin Heidelberg New York 1995) 24
265. Kennicutt, R. C., Jr. 1998, Ann. Rev. Astron. Astrophys. **36**, 189 51, 188, 189
266. Kereš, D., Katz, N., Weinberg, D. H., Davé, R. 2005, MNRAS **363**, 2 289
267. Kim, T. S., Christiani, S., D'Odorico, S. 2001, A&A **373**, 757 200, 201
268. Kippenhahn, R., Weigert, A. *Stellar Structure and Evolution* (Springer, Berlin Heidelberg, New York 1990) 17, 20, 22, 54
269. Klessen, R. S., Spaans, M., Jappsen, A.-K. 2007, MNRAS **374**, L29 317
270. Kobulnicky, H. A., Kewly, L. J. 2004, ApJ **617**, 240 191
271. Kobulnicky, H. A., Zaritsky, D. 1999, ApJ **511**, 188 42
272. Kodama, T., Tanaka, I., Kajisawa, M., et al. 2007 MNRAS **377**, 1717 276
273. Köppen, J., Weidner. C., Kroupa, P. 2007, MNRAS **375**, 673 297
274. Kormendy, J. 1977, ApJ **218**, 333 35
275. Kormendy, J., Bender, R. 1996, ApJ **464**, L119 33
276. Kormendy, J., Richstone, D.O. 1995, Ann. Rev. Astron. Astrophys. **33**, 518 56
277. Korn, A. J., et al. 2006, Nature **442**, 657 82
278. Kotilainen, J. K., Falomo, R., Labita, M., et al. 2007, ApJ **660**, 1039 264
279. Kovač, K., Somerville, R. S., Rhoads, J. E., et al. 2007, ApJ **668**, 415 272, 274
280. Kreysa, E., et al. 1998, SPIE **3357**, 319 135
281. Kriek, M., van Dokkum, P. G., Franx, M., et al. 2006, ApJ **645**, 44 223, 224, 259
282. Kriek, M., van Dokkum, P. G., Franx, M., et al. 2006, ApJ **649**, L71 258, 259
283. Kriek, M., van Dokkum, P. G., Franx, M., et al. 2007, ApJ **669**, 776 264
284. Kriek, M., van der Wel, A., van Dokkum, P. G., et al. 2008, ApJ **682**, 896 259, 261, 263
285. Kroupa, P. 2001, MNRAS **322**, 231 13
286. Kroupa, P. 2007, Proc. IAU Symposium **241** (ed. A. Vazdekis and R. F. Peletier) p. 109 13

287. Krügel, E. *The Physics of Interstellar Dust* (Institute of Physics, Bristol, UK 2003) 49, 150, 199
288. Kudritzki, R.-P., Méndez, R. H., Feldmeier, J. J., et al. 2000, ApJ, 536, 19 248
289. Kukula, M.J., Dunlop, J. S., McLure, R. J., et al. 2001, MNRAS **326**, 1533 264
290. Kunth, D., Mas-Hesse, J. M., Terlevich, E., et al. 1998, A&A **334**, 11 254
291. Kunth, D., Leitherer, Mas-Hesse, J. M., et al. 2003, ApJ **597**, 263 51, 53, 253
292. Kurk, J. D., Walter, F., Fan, X., et al. 2007, ApJ **669**, 32 264, 265
293. Kuster, M., Raffelt, G., Beltr'an, B. (eds.) *Axions. Theory, Cosmology, and Experimental Searches*, Springer Lecture Notes in Physics Vol. **741** (Springer, Berlin Heidelberg, New York 2008) 70
294. Labbé, I., Franx, M., Rudnick, G., et al. 2003, AJ **125**, 1107 170
295. Labbé, I., Huang, J., Franx, M., et al. 2005, ApJ **624**, 81 258, 263
296. Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2008 MNRAS **385**, 1155 317
297. Lai, K., Huang, J.-S., Fazio, G., et al. 2007, ApJ **655**, 704 251
298. Lai, K., Huang, J.-S., Fazio, G., et al. 2008, ApJ **674**, 70 251
299. Landy, S. D., Szalay, A. S. 1993, ApJ **412**, 64 271
300. Lapi, A., Kawakatu, N., Bosnjak, Z., et al. 2008, MNRAS **386**, 608 267, 304
301. Lauer, T. R., Gebhardt, K., Faber, S. M., et al. 2007, ApJ **664**, 226 56
302. Law, D. R., Steidel, C. C., Erb, D. K., et al. 2007, ApJ **669**, 929 223, 240
303. Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS **379**, 1599 165
304. Leavitt, H. S., 1908, Harvard Obs. Ann. **60**, 87 7
305. Ledoux, P. 1941, ApJ **94**, 537 20
306. Le Fèvre, O., Mellier, Y, McCracken, H. J., et al. 2004, A&A **417**, 839 163, 166
307. Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A **439**, 845 166, 208
308. Leibundgut, B. 2001, Ann. Rev. Astron. Astrophys. **39**, 67 28
309. Leitherer, C. 2008, Proc. IAU Symp. **255**, 305 316
310. Leitherer, C., Leao, J. R. S., Heckman, T. M., et al. 2001, ApJ **550**, 724 183
311. Leitherer, C., Li, I.-H., Calzetti, D., et al. 2002, ApJS **140**, 303 53, 181
312. Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, **123**, 3 181, 183, 184, 185, 186
313. Léna, P., Lebrun, F., Mignard, F. *Observational Astrophysics*, 2nd edn (Springer, Berlin Heidelberg, New York 1998) 129
314. Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS **172**, 70 166, 208
315. Limber, D. N. 1953, ApJ **117**, 134 271
316. Lin, L., Patton, D. R., Koo, D. C., et al. 2008, ApJ **681**, 232 305
317. Lin, C. C., Shu, F. H. 1964, ApJ **140**, 646 36
318. Liu, X., Shapley, A. E., Coil, A. L., et al. 2008, ApJ **678**, 758 42, 235
319. Longair, M.: *Galaxy Formation*, 2nd edn (Springer, Berlin Heidelberg, New York 2008) 72, 78, 88, 90, 91
320. Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, PASP **115**, 897 165
321. Lotz, J. M., Madau, P., Giavalisco, M., et al. 2006, ApJ **636**, 592 212
322. Lundmark, K. 1919, Astr. Nachrichten **209**, 360 7
323. Lutz, D., Sturm, E., Tacconi, L. J., et al. 2008, ApJ **684**, 853 265
324. Lutz, D., Valiante, E., Sturm, E., et al. 2005, ApJ **625**, L83 256, 257, 304
325. MacDonald, E. C., Allen, P., Dalton, D., et al. 2004, MNRAS **352**, 1255 166
326. Madau, P., Ferarra, A., Rees, M. J. 2001, ApJ **555**, 92 287
327. Madau, P., Pozzetti, L., Dickinson, M. 1998, ApJ **498**, 106 219
328. Madau, P., Rees, M. J. 2001, ApJ **551**, L27 303
329. Maiolino, R., Cox, P., Caselli, P., et al. 2005, A&A **440**, L51 265
330. Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A **488**, 463 223, 230, 231, 296, 297
331. Maiolino, R., Neri, R., Beelen, F., et al. 2007, A&A **472**, L33 264
332. Maiolino, R., Schneider, R., Oliva, E., et al. 2004, Nature **431**, 533 265
333. Malhotra, S., Rhoads, J. E. 2004, ApJ **617**, L5 248
334. Malhotra, S., Rhoads, J. E., Pirzkal, N., et al. 2005, ApJ **626**, 666 208, 269
335. Mansouri, R., Brandenberger, R., (eds) *Large Scale Structure Formation* (Kluwer, Dordrecht 2000) 88
336. Maraston, C. 2005, MNRAS **362**, 799 183, 260

337. Maraston, C., Daddi, E., Renzini, A., et al. 2006, ApJ **652**, 85 317
338. Martin, D. C., Fanson, J., Schiminovich, D., et al. 2005, ApJ **619**, L1 165
339. Mas-Hesse, J.M. , Kunth, D., Tenorio-Tagle, G., et al. 2003, ApJ **598**, 858 240, 254
340. Mather, J. C., Fixsen, D. J., Shafer, R. A., et al. 1999, ApJ **512**, 511 69
341. Matsuda, Y., Yamada, T., Hayashinao, T., et al. 2004, AJ **128**, 569 255
342. McKee, C. F., Tan, J. C. 2008, ApJ **681**, 771 114
343. McMahon, R., Parry, I., Venemans, B., et al. 2008, The Messenger **131**, 11 313
344. Meadows, A. *The Future of the Universe* (Springer, Berlin Heidelberg, New York 2006) 116
345. Mehlert, D., Noll, S., Appenzeller, I., et al. 2002, A&A **393**, 809 194, 232
346. Mehlert, D., Seitz, S., Saglia, R. P., et al. 2001, A&A **379**, 96 171, 172
347. Mehlert, D., Tapken, C., Appenzeller, I., et al. 2006, A&A **455**, 835 195, 196, 223, 233, 237, 238, 239, 240
348. Meinköhn, E., & Richling, S. 2002, A&A **392**, 827 254
349. Mészáros, P. 2002, Ann. Rev. Astron. Astrophys. **40**, 137 26
350. Metcalfe, N., Shanks, T., Campos, A., et al. 2001, MNRAS **323**, 795 168
351. Meuerer,G. R., Heckman, T. M., Calzetti, D. 1999, ApJ **521**, 64 181, 242
352. Meuerer, G. R., Tsvetanov, Z.I., Gronwall, C., et al. 2007, AJ **134**, 77 159
353. Mignoli, M., Cimatti, A., Zamorani, G., et al. 2005, A&A **437**, 883 170
354. Miley, G. K., Overzier, R. A., Tsvetanov, Z. I., et al. 2004, Nature **427**, 47 274
355. Minkowski, R. 1960, ApJ **132**, 908 10
356. Miralda-Escude, J., Rees, M. J. 1997, ApJ **478**, L57 288
357. Misner, C. W., Thorne, K. S., Wheeler, J. A. *Gravitation* (W. H. Freeman and Company, San Francisco, 1973) Chapter 33. 29
358. Mo, H. J., Mao, S., White, S. D. M. 1998, MNRAS **295**, 319 110, 112, 294
359. Mo, H. J., Mao, S., White, S. D. M. 1999, MNRAS **304**, 175 110
360. Mobasher, B., Capak, P., Scoville, N. Z., et al. 2007, ApJS **172**, 117 178
361. Moles, M., Benítez, N., Aguerri, J. A. L., et al. 2008, AJ **136**, 1325 166
362. Møller, P., Warren S. J., Fall, S. M. 2002, A&A **574**, 51 291
363. Morgan, W. W., Abt, H. A., Tapscott, J. W. *Revised MK Atlas for Stars Earlier than the Sun,* Kitt Peak National Observatory, Tucson, Arizona, 1978 15
364. Mortier, A. M. J., et al. 2005, MNRAS **363**, 563 170
365. Moustakas, L. A., Somerville, R. S. 2002, ApJ **577**, 1 274
366. Murayama, T., Taniguchi, Y., Scoville, N. Z., et al. 2007, ApJS **172**, 523 209, 248, 251
367. Nagao, T., Maiolino, R., Marconi, A. 2006, A&A **459**, 85 191, 192
368. Nagao, T., Sasaki, S. S., Maiolino, R., et al. 2008, ApJ **680**, 100 300
369. Naselsky, P. D., Novikov, D. I., Novikov, I. D. *The Physics of the Cosmic Microwave Background* (Cambridge University Press, Cambridge, UK, 2006) 88, 100
370. Navarro, J. F., Frenk, C. S., White, S. D. M. 1996, ApJ **462**, 535 97
371. Nesvadba, N. P. H., Lehnert, M. D., Davies, R. I., et al. 2008, A&A **479**, 67 240
372. Nesvadba, N. P. H., Lehnert, M. D., Eisenhauer, F., et al. 2006, ApJ **650**, 661 242, 305
373. Neufeld, D. A. 1991, ApJ **370**, L85 253
374. Ni, K., Baudis, L. 2008, Proc. 36th COSPAR Scientific Assembly, Adv. Space Res. **41**, 2019 70
375. Nilsson, K. K., Fynbo, J. P. U., Møller, P., et al. 2006, A&A **452**, L23 255
376. Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ **660**, 43 304
377. Noll, S., Mehlert, D., Appenzeller, I., et al. 2004, A&A **418**, 885 154, 207, 222, 227, 228, 229, 230, 243
378. Noll, S., Pierini, D. 2005, A&A **444**, 137 182, 198, 244, 245, 246
379. Noll, S., Pierini, D., Pannella, M., Savaglio, S. 2007, A&A **472**, 455 182, 198, 245
380. Nozawa, T., Kozasa, T., Umeda, H., Maeda, K., Nomoto, K. 2003, ApJ **598**, 785 301
381. Odewahn, S. C., de Carvalho, R. R., Gall, R. R., et al. 2005, AJ **128**, 3092 164
382. Oesch, P. A., Stiaelli, M, Carollo, C. M., et al. 2007, ApJ **671**, 1212 167, 208
383. Oke, J. B. 1965, Ann. Rev. Astron. Astrophys. **3**, 23 140
384. Oke, J. B., Gunn, J. E. 1983, ApJ **266**, 713 140
385. Omont, A., Cox, P., Bertoldi, F., et al. 2001, A&A **374**, 371 199
386. Omukai, K., Inutsuka, S. 2002, MNRAS **332**, 59 114
387. Omukai, K., Nishi, R. 1998, ApJ **508**, 141 114

388. Oort, J. H. 1940, ApJ **91**, 273 69
389. Öpik, E. 1922, ApJ **55**, 406 7
390. Osterbrock, D. E. *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei* (University Science Books, Mill Valley 1989) 190, 191, 197
391. Osterbrock, D., E., Fulbright, J. P., Martel, A. R., et al. 1996, PASP **108**, 227 145
392. Ostriker, J. P., Peebles, P. J. E. 1973, ApJ, **186**, 467 37
393. Ota, K., Iye, M., Kashikawa, N., et al. 2008, ApJ **677**, 12 286
394. Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2008, ApJS **176**, 301 208, 248
395. Ouchi, M., Shimasaku, K. Okamura, S. 2004, ApJ **611**, 685 207
396. Ouchi, M., Shimasaku, K, Akiyama, M. et al. 2005, ApJ **620**, 1 272, 274
397. Overzier, R. A., Bouwens, R. J., Cross, N. J. G., et al. 2008, ApJ **673**, 143 274, 275
398. Overzier, R. A., Heckman, T. M., Kauffmann, G., et al. 2008, ApJ **677**, 37 306
399. Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., et al. 1979, MNRAS **189**, 95 191
400. Papovich, C., Dickinson, M., Giavalisco, M. 2005, ApJ **631**, 101 176, 212
401. Partridge, R. B., Peebles, P. J. E. 1967, ApJ, **147**, 868 158, 205, 246
402. Pasquali, A., Pirzkal, N., Larsen, S., et al. 2006, PASP **118**, 270 147
403. Peacock, J. A. *Cosmological Physics* (Cambridge University Press 1999) 73
404. Peebles, P. J. E. *The Large Scale Structure of the Universe* (Princeton University Press 1980) 108
405. Peebles, P. J. E. *Principles of Physical Cosmology* (Princeton University Press 1993) 73, 84, 86
406. Peebles, P. J. E., Ratra, B., 1988, ApJ **325**, L17 71
407. Pelló, R. Probing Distant Galaxies with Lensing Clusters. In: J. Bergeron, A. Renzini (eds) *From Extrasolar Planets to Cosmology: The VLT Opening Symposium* (Springer, Berlin Heidelberg, New York 2000) p. 131 171
408. Penzias, A. A., Wilson, R. W. 1965, ApJ **142**, 419 4
409. Percival, W. J., Nichol, R. C., Eisenstein, D. J. 2007, ApJ **657**, 645 104, 105
410. Perez-Gonzalez, P. G., Trujillo, I., Barro, G., et al. 2008, ApJ **687**, 50 260
411. Petrosian, V. 1976, ApJ **209**, L1 38
412. Pettini, M., Ellison, S. L., Bergeron, J., Petitjean, P. 2002, A&A **391**, 21 290
413. Pettini, M., Ellison, S. L., Steidel, C. C., et al. 2000, ApJ **532**, 65 290
414. Pettini, M., Madau, P., Bolte, M., et al. 2003, ApJ **594**, 695 289
415. Pettini, M., Rix, S. A., Steidel, C. C., et al. 2002, ApJ **569**, 742 234, 238, 239
416. Pettini, M., Steidel, C. C., Adelberger, K. L., et al. 2000, ApJ **528**, 96 224, 234, 236, 238
417. Pettini, M., Zych, B. J., Steidel, C. C., Chaffee, F. H. 2008, MNRAS **385**, 2011 291
418. Piacentini, F., Ade, P. A. R., Bhatia, R. S. 2002, ApJS **138**, 315 133
419. Pickering, E. C. 1912, Harvard Circ. No. 173 7
420. Pirzkal, N., Malhotra, S., Rhoads, J. E., Xu, C. 2007, ApJ **667**, 49 251
421. Press, W. H., Schechter, P., 1974, ApJ **187**, 425 109
422. Prochaska, J. X., Chen, H.-W., Dessauges-Zavadsky, M., et al. 2007, ApJ **666**, 267 267
423. Quadri, R., van Dokkum, P. Gawiser, E., et al. 2007, ApJ **654**, 138 272, 273
424. Rauch, M. 1998, Ann. Rev. Astron. Astrophys. **36**, 267 107, 278
425. Rauch, M., Haehnelt, M. G., Bunker, A., et al. 2008 ApJ **681**, 856 291
426. Rauch, M., Haehnelt, M. G., Steinmetz, M. 1997, ApJ **481**, 601 288
427. Reddy, N. A., Erb, D. K., Steidel, C. C., et al. 2005, ApJ **633**, 748 256, 258
428. Reddy, N. A., Steidel, C. C., Erb, D. K., et al. 2006, ApJ **653**, 1004 208
429. Reddy, N. A., Steidel, C. C., Pettini, M., et al. 2008, ApJS **175**, 48 207, 219, 249
430. Reid, I. N., Brewer, C., Brucato, R. J., et al. 1991, PASP **103**, 661 164
431. Renzini, A. 2006, Ann. Rev. Astron. Astrophys. **44**, 141 35, 186, 196
432. Rhoads, J. E., Malhotra, S. 2001, ApJ **563**, L5 248
433. Rhook, K. J., Haehnelt, M. G. 2006, MNRAS **373**, 623 324
434. Ridgway, S. E., Heckman, T. M., Calzetti, D., Lehnert, M. 2001, ApJ **550**, 122 264
435. Rieke, G. H. 2007, Ann. Rev. Astron. Astrophys. **45**, 77 134
436. Riess, A. G., et al. 2004, ApJ **607**, 665 71, 116
437. Rigaut, F. J., Ellerbroek, B. L., Flicker, R. 2000, SPIE **4007**, 1022 126
438. Rix, H.-W., Barden, M., Beckwith, S. V. W., et al. 2004, ApJS **152**, 163 166

439. Rix, S. A., Pettini, M., Leitherer, C., et al. 2004, ApJ **615**, 98 183, 193, 195, 224
440. Roddier, F. *Adaptive Optics in Astronomy* (Cambridge University Press, Cambridge 1999) 126
441. Rohlfs, K., Wilson, T. L.: *Tools of Radio Astronomy*, 4nd printing (Springer, Berlin Heidelberg, New York 2004) 124, 131, 146, 149, 197
442. Sawicki, M., Thompson, D. 2005, ApJ **635**, 100 168
443. Saito, T., Shimasaku, K., Okamura, S., et al. 2006, ApJ **648**, 54 255
444. Salimbeni, S., Gaillongo, E., Menci, N., et al. 2008, A&A **477**, 763 261, 262
445. Salpeter, E. E. 1955, ApJ **121**, 161 13
446. Sandage, A. 1988, Ann. Rev. Astron. Astrophys. **26**, 561 5
447. Sandage, A., Tammann, G. A. 2006, Ann. Rev. Astron. Astrophys. **44**, 93 7
448. Sandage, A., Tammann, G. A., Yahil, A. 1979, ApJ **232**, 352 183
449. Santos, M. R., Ellis, R. S., Kneib, J.-P., Richard, J., Kuijken, K. 2004, ApJ **606**, 683 171
450. Savaglio, S., Glazebrook, K., Le Borgne, D., et al. 2005, ApJ **635**, 260 231
451. Scarlata, C., Carollo, C. M., Lilly, S. J., et al. 2007, ApJS **172**, 494 305
452. Schaerer, D. 2003, A&A **397**, 527 225, 249, 299, 300
453. Schaerer, D., Verhamme, A. 2008, A&A **480**, 369 227, 236
454. Schawinski, K., Thomas, D., Sarzi, M., et al. 2007, MNRAS **382**, 1415 302, 305
455. Schechter, P. 1976, ApJ **203**, 297 39
456. Schmidt, M. 1959, ApJ **129**, 243 51
457. Schneider, P. *Extragalactic Astronomy and Cosmology* (Springer, Berlin Heidelberg, New York 2006) 32, 73
458. Schneider, P., Ehlers, J., Falco, E. E. *Gravitational Lenses*, 2nd printing (Springer, Berlin Heidelberg, New York 1999) 65, 242
459. Schneider, D., Gun, J., Turner, E., et al. 1986, AJ **91**, 991 10
460. Schödel, R., Ott, T., Genzel, R., et al. 2003, ApJ **596**, 1015 56
461. Schramm, M., Wisotzki, L., Jahnke, K. 2008 A&A **478**, 311 264
462. Schwarzschild, M., Härm, R., 1959, ApJ **129**, 637 20
463. Schweitzer, M., Lutz, D., Sturm, E., et al. 2006, ApJ **649**, 79 256
464. Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS **172**, 1 169
465. Sekiguchi, K., Akiyama, M., Furusawa, H., et al. In: A. Renzini and R.Bender (eds) *Multiwavelength Mapping of Galaxy Formation and Evolution* (Springer, Berlin Heidelberg NewYork 2005) p. 82 169
466. Seldner, M., Siebers, B., Groth, E. J., Peebles, P. J. E. 1977, AJ **82**, 249 64, 269
467. Sérsic, J. L. 1968: *Atlas de Galaxias Australes* (Observatorio Astronomico de Cordoba, Cordoba, Argentina, 1968) 33
468. Seymour, N., Ogle, P., De Breuck, C., et al. 2008, ApJ **681**, L1 266
469. Seymour, N., Stern, D., De Breuck, C., et al. 2007, ApJS **171**, 353 266, 296
470. Shapley, A. E., Steidel, C. C., Erb, D. K., et al. 2005, ApJ **626**, 698 224
471. Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L. 2003, ApJ **588**, 65 181, 222, 226, 228, 238, 239
472. Shapley, A. E., Steidel, C. C., Pettini, M., et al. 2006, ApJ **651**, 688 185, 283, 284
473. Shapley, H. 1917, PASP **29**, 213 7
474. Shen, S., Mo, H. J., White, S. D.M., et al. 2003, MNRAS **343**, 978 38, 295
475. Shimasaku, K., Kashikawa, N., Doi, M., et al. 2006, PASJ **58**, 313 209, 248, 249
476. Silk, J. 1968, ApJ **151**, 459 91
477. Simcoe, R. A. 2006, ApJ **653**, 977 288, 289
478. Siringo, G., Weiss, A., Kreysa, E., et al. 2007, The Messenger **129**, 2 135
479. Skrutskie, M. F., Cutrie, R. M., Stiening, R., et al. 2006, AJ**131**, 1163 164
480. Slipher, V. M. 1915, Popular Astronomy **23**, 21 8
481. Smail, I., Swinbank, A. M., Richard, J., et al. 2007, ApJ **654**, 33 171
482. Smith, J.A., Tucker, D. L., Kent, S., et al. 2002, AJ **123**, 2121 141
483. Sofue, Y., Rubin, V. 2001, Ann. Rev. Astron. Astrophys. **39**, 137 69
484. Somerville, R. S., Primack, J. R. 1999, MNRAS **310**, 1087 97
485. Songaila, A. 2001, ApJ, **561**, L153 278, 288, 289
486. Songaila, A. 2004, AJ **127**, 2598 279
487. Songaila, A. 2005, AJ **130**, 1996 288

488. Songaila, A., Cowie, L. L. 1996, AJ **112**, 335 288
489. Soifer, B. T., Neugebauer, G., Houck, J. R. 1987, Ann. Rev. Astron. Astrophys. **25**, 187 164
490. Spaans, M., Silk, J. 2006, ApJ **652**, 902 303
491. Spergel, D. N., Bean, R., Doré, O., et al. 2007, ApJS **170**, 377 3, 70, 71, 76, 100, 103, 116, 285
492. Springel, V., White, S. D. M., Jenkins, A. 2005, Nature **435**, 629 94, 95, 97, 98, 109, 293, 296, 298, 304
493. Stanway, E. R., Bunker, A. J., McMahon, R. G., 2004, ApJ **607**, 704 249
494. Stark, D. P., Ellis, R. S., Richard, J., et al. 2007, ApJ **633**, 10 216
495. Steidel, C. C., Adelberger, K. L., Giavalisco, M., et al. 1998, ApJ **492**, 428 269
496. Steidel, C. C., Adelberger, K. L., Giavalisco, M., et al. 1999, ApJ **519**, 1 156, 214, 215
497. Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2000, ApJ **532**, 170 255
498. Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, ApJ **592**, 728 156, 207, 269
499. Steidel, C. C., Dickinson, M., Sargent, W. L. 1991, AJ **101**, 1187 10
500. Steidel, C. C., Giavalisco, M., Dickinson, M., et al. 1996, AJ **112**, 352 156
501. Steidel, C. C., Giavalisco, M., Pettini, M., et al. 1996, ApJ **462**, L17 305
502. Steidel, C. C., Hamilton, D., 1992, AJ **104**, L93 156
503. Stratta, G., Maiolino, R., Fiore, F., et al. 2007, ApJ **661**, L9 267
504. Stutz, A. M., Papovich, C., Eisenstein, D. J. 2008, ApJ **677**, 828 263
505. Tacconi, L. J., Neri, R., Chapman, S. C., et al. 2006, ApJ **640**, 228 258, 304
506. Tacconi, L. J., Genzel, R., Smail, I., et al. 2008, ApJ **680**, 246 258
507. Takata, T., Sekiguchi, K., Smail, I., et al. 2006, ApJ **651**, 713 256
508. Tammann, G. A. 2006, Rev. Modern Astronomy **19**, 1 9
509. Tan, J. C., McKee, C. F. 2004, ApJ **603**, 383 114
510. Tanaka, M., Kodama, T., Arimoto, N., et al. 2005, MNRAS **362**, 268 45, 46
511. Tapken, C., Appenzeller, I., Gabasch, A., et al. 2006, A&A **455**, 145 248
512. Tapken, C., Appenzeller, I., Gabasch, A., et al. 2007, The Messenger **128**, 51 159, 161
513. Tapken, C., Appenzeller, I., Mehlert, D., et al. 2004, A&A **416**, L1 224, 247, 250, 252, 254
514. Tapken, C., Appenzeller, I., Noll, S., et al. 2007, A&A **467**, 63 249, 250, 251, 254
515. Tenorio-Tagle, G., Muñoz-Tuñón, Pérez, E., et al. 2006, ApJ **643**, 186 240, 254
516. Tenorio-Tagle, G., Silich, S., Muñoz-Tuñón, et al. 2003, ApJ **597**, 279 240
517. Theuns, T., Viel, M., Kay, S., et al. 2002, ApJ **578**, L5 287
518. Thommes, E. , Meisenheimer, K. 2005, A&A, 430, 877 251
519. Thompson, L. A., Gardner, C. S. 1987, Nature **328**, 229 126
520. Thompson, R. I., Illingworth, G., Bouwens, R., et al. 2005, AJ **130**, 1 167
521. Toft, S., van Dokkum, P., Franx, M., et al. 2007, ApJ, **671**, 285 214
522. Treister, E., Gawiser, E., Van Dokkum, P., et al. 2007, The Messenger, **129**, 45 166
523. Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ **613**, 898 42, 43, 231
524. Tresse, L., Ilbert, O., Zucca, E., et al. 2007, A&A **472**, 403 219
525. Trimble, V. 1996, PASP **108**, 1073 9
526. Trujillo, I., Feulner, G., Goranova, Y., et al. 2006, MNRAS **373**, L36 263
527. Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, ApJ **650**, 18 213
528. Trümper, J., 1993, Science **260**, 1769 164
529. Tully, R. B., Fisher, J. R. 1977, A&A **54**, 661 37
530. Tumlinson, J. 2003, ApJ **584**, 608 300
531. Tumlinson, J. 2007, ApJ **665**, 1361 301
532. Ulmschneider, P. *Intelligent Life in the Universe* (Springer, Berlin Heidelberg, New York 2003) 12
533. Umeda, H., Nomoto, K. 2003, Nature **422**, 871 301
534. Väisänen, P., Mattila, S., Kniazev, A., et al. 2008, MNRAS **384**, 886 53
535. van Breukelen, C., Jarvis, M. J., Venemans, B. P. 2005, MNRAS **359**, 895 248
536. van den Bosch, F. C., Yang, X., Mo, H. J., et al. 2003, MNRAS **340**, 771 271
537. van Dokkum, P. G. 2008, ApJ **674**, 29 317
538. van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, ApJ **677**, L5 263
539. van Dokkum, P. G., Marchesini, D., Labbé, et al. 2008, NOAO/NSO Newsletter, June 2008, p. 8 313
540. van Waerbeke, L., Mellier, Y., Hoekstra, H. 2005, A&A **429**, 75 107

541. Vanzella, E., Cristiani, S., Dickinson, M., et al. 2006, A&A **454**, 423 229, 249
542. Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, A&A **478**, 83 208, 250
543. Vazquez, G. A., Leitherer, C. 2005, ApJ **621**, 695 183
544. Venemans, B. P., Roettgering, H. J. A., Miley, G. K. 2005, A&A **431**, 793 250
545. Verhamme, A., Schaerer, D., Maselli, A. 2006, A&A **460**, 397 254
546. Verhamme, A., Schaerer, D., Atek, H., Tapken, C. 2008, A&A **491**, 89 254
547. Viel, M., Matarrese, S., Mo, H. J., et al. 2002, MNRAS **336**, 685 279
548. Villforth, C., Heidt, J., Nilsson, K. 2008 A&A **488**, 133 264
549. Walborn, N. R., Ebbets, D. C., Parker, J. W., et al. 1992, ApJ **393**, L13 16
550. Walter, F., Carilli, C., Bertoldi, F., et al. 2004, ApJ **615**, L17 264, 296
551. Wambsganss, J., Ostriker, J. P., Bode, P. 2008, ApJ **676**, 753 108
552. Wang, R., Wagg, J., Carilli, C. L., et al. 2008, AJ **135**, 1201 264, 265
553. Web, T.M., Eales, S., Foucaud, S., et al. 2003, ApJ **582**, 6 170
554. Westra, E., Heath Jones, D., Lidman, C. E., et al. 2006, A&A **455**, 61 248
555. Wetterich, C. 1988, Nucl. Phys. **B302**, 668 71
556. Wielebinski, R., Beck, R. *Cosmic Magnetic Fields* (Springer, Berlin Heidelberg, New York 2005) 49
557. Williams, R. E., Baum, S., Bergeron, L. E., et al. 2000, AJ **120**, 2735 167
558. Williams, R. E., Blacker, B., Dickinson, M., et al. 1996, AJ **112**, 1335 167
559. Wilson, R. N. *Reflecting Telescope Optics I*, 2nd edn (Springer, Berlin Heidelberg, New York 2004) 123
560. Wilson, R. N. *Reflecting Telescope Optics II* (Springer, Berlin Heidelberg New York 1999) 123, 124
561. Wirtz, C. 1922, Astronomische Nachrichten **215**, 349 8
562. Wirtz, C. 1924, Astronomische Nachrichten **222**, 21 8
563. Wolf, C., Meisenheimer, K., Rix, H.-W., et al. 2003, A&A **401**, 73 165, 178
564. Wolf, C., Meisenheimer, K., Kleinheinrich, M., et al. 2004, A&A **421**, 913 166
565. Wolfe, A. M., Gawiser, E., Prochaska, J. X. 2005, Ann. Rev. Astron. Astrophys. **43**, 861 290, 291
566. Woosley, S. E., Bloom, J. S. 2006, Ann. Rev. Astron. Astrophys. 44, 507 25
567. Wyder, T. K., Treyer, M. A., Milliard, B. et al. 2005, ApJ **619**, L15 215
568. Xu, C., Pirzkal, N., Malhotra, S., et al. 2007, AJ **134**, 169 159
569. York, D. G., Adelman, J., Anderson, J. E., et al. 2000, AJ **120**, 1579 164
570. Yoshida, N., Abel, T., Hernquist, L., Sugiyama, N. 2003, ApJ **592**, 645 113
571. Yoshida, N., Omukai, K., Hernquist, L. 2007, ApJ **667**, L117 113
572. Yoshida, M., Shimasaku, K., Kashikawa, N., et al. 2006, ApJ **653**, 988 156, 157, 207
573. Yoshida, M., Shimasaku, K., Ouchi, M., et al. 2008, ApJ **679**, 269 272
574. Zaldarriaga, M. 1997, Phys. Rev. D **55**, 1830 285
575. Zaldarriaga, M., 2000, An Introduction to CMB An Isotropies. In: R. Mansouri, R. Brandenberger (eds) *Large Scale Structure Formation* (Kluwer, Dordrecht 2000), p. 213 89, 100
576. Zaldarriaga, M., Seljak, U. 2000, ApJS **129**, 431 89
577. Zehavi, I., Weinberg, D. H., Zheng, Z., et al. 2004 ApJ **608**, 16 271
578. Zheng, W., Overzier, R. A., Bouwens, R. J., et al. 2006, ApJ **640**, 574 274, 275
579. Zinnecker, H., Yorke, H. W. 2007, Ann. Rev. Astron. Astrophys. **45**, 481 20
580. Zwicky, F. 1933, Helv. Phys. Acta **6**, 110 65, 69

Index

ASTRONOMY AND ASTROPHYSICS LIBRARY

ASTRONOMY AND ASTROPHYSICS LIBRARY

The Design and Construction of Large Optical Telescopes By P. Y. Bely

The Solar System 4th Edition
By T. Encrenaz, J.-P. Bibring, M. Blanc, M. A. Barucci, F. Roques, Ph. Zarka

General Relativity, Astrophysics, and Cosmology By A. K. Raychaudhuri, S. Banerji, and A. Banerjee

Stellar Interiors Physical Principles, Structure, and Evolution 2nd Edition
By C. J. Hansen, S. D. Kawaler, and V. Trimble

Asymptotic Giant Branch Stars
By H. J. Habing and H. Olofsson

The Interstellar Medium
By J. Lequeux

Methods of Celestial Mechanics (2 volumes)
Volume I: Physical, Mathematical, and Numerical Principles
Volume II: Application to Planetary System, Geodynamics and Satellite Geodesy
By G. Beutler

Solar-Type Activity in Main-Sequence Stars
By R. E. Gershberg

Relativistic Astrophysics and Cosmology
A Primer By P. Hoyng

Magneto-Fluid Dynamics
Fundamentals and Case Studies
By P. Lorrain

Compact Objects in Astrophysics
White Dwarfs, Neutron Stars and Black Holes
By Max Camenzind

Special and General Relativity
With Applications to White Dwarfs, Neutron Stars and Black Holes
By Norman K. Glendenning

Planetary Systems
Detection, Formation and Habitability of Extrasolar Planets
By M. Ollivier, T. Encrenaz, F. Roques F. Selsis and F. Casoli

The Sun from Space 2nd Edition
By Kenneth R. Lang

Tools of Radio Astronomy 5th Edition
By Thomas L. Wilson, Kristen Rohlfs and Susanne Hüttemeister

Astronomical Optics and Elasticity Theory
Active Optics Methods
By Gérard René Lemaitre

High-Redshift Galaxies
Light from the Early Universe
By I. Appenzeller