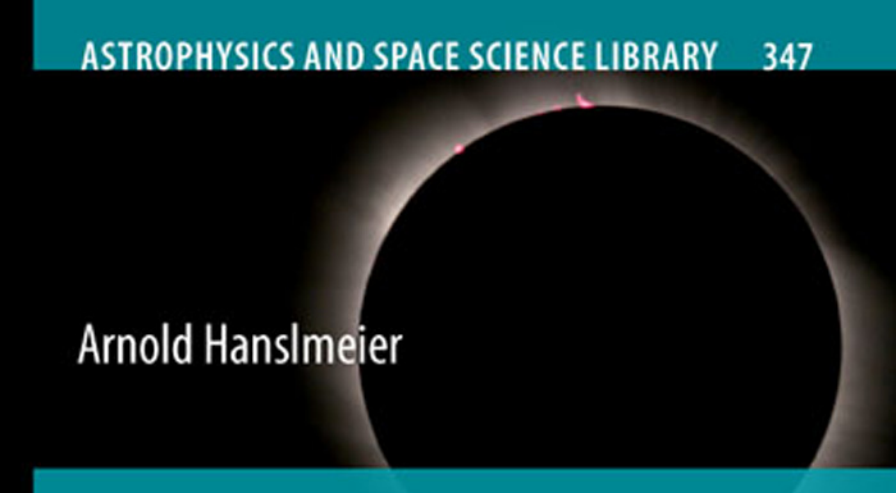
ASTROPHYSICS AND SPACE SCIENCE LIBRARY 347



Arnold Hanslmeier

The Sun and Space Weather

2nd Edition

ASSL

Springer

ASTROPHYSICS AND SPACE SCIENCE LIBRARY

VOLUME 347

THE SUN AND SPACE WEATHER

Second Edition

by

ARNOLD HANSLMEIER
University of Graz, Institute of Physics/ IGAM, Austria

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-5603-6 (HB)
ISBN-13 978-1-4020-5603-1 (HB)
ISBN-10 1-4020-5604-4 (e-book)
ISBN-13 978-1-4020-5604-8 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Photo cover: Solar Elipse 2006 by Prof. Dr. A. Hanslmeier
Printed on acid-free paper

I want to thank my students who attended a course on space weather I held at Graz and Innsbruck university; they critically read the manuscript and suggested corrections. I also want to thank my colleagues who contributed Figures and gave many hints. Special thanks to my wife Caroline and my children Roland, Christina and Alina; I was allowed to spend lots of nights at the PC. Thanks for the patience of my collaborators.

Contents

Preface

The field of solar physics and solar--terrestrial relation, now called space weather, is evolving rapidly. As in the first edition, it is assumed that it is inevitable for the reader to get some basic knowledge in solar physics since the Sun is the main driver for space weather The term *space weather* itself has been gaining more and more attention during the past years as our society becomes more and more dependent on satellites, which are vulnerable to varying conditions in space. Space weather efforts and investigations are being made all over the world and more and more is known about the complex relations of processes on the Sun and the Earth and its space environment. The term *space climate* nowadays includes the long-term variations caused mainly by the Sun on the Earth and the interplanetary space.

As in the first edition of the book, this edition also covers these topics but new chapters have been introduced, e.g., a chapter on real-time space weather forecasts and some main space weather data sources. All the chapters have updated information, taking into account the results of new satellite missions and telescopes. The book also includes a great amount of new literature (more than 340 original citations) so that the reader is able to go into more details, if required in the respective chapters. Thus, the book should be helpful to scientists as well as to students interested in overview or finding a compendium with references to go deeper into special fields.

Furthermore, at the beginning of all the chapters, introductory books are cited, which could be recommended for the special topics addressed there. The number of keywords in the index has also been strongly enhanced so that the reader can find information easily. Besides all this, suggestions from readers of the first edition have been taken into account and are greatly acknowledged.

I want to thank all my colleagues who provided me with advice and figures and the students who attended my lectures at Graz and Innsbruck for their help. Last but not least I thank my family – Karoline, Roland, Christina and Alina – for the patience and understanding when I spent lots of nights at the computer.

Chapter 1

Introduction, What is Space Weather?

In this introduction we briefly describe the term space weather and give motivation why that interdisciplinary field gained high interest. Examples will demonstrate the high relevance of space weather not only from the scientific point of view but from the social and economic aspect of our modern civilization.

Since this is a very modern topic, there appeared several monographs about that subject, e.g. a collection of space weather related topics[1].

1.1 Definition of Space Weather

Modern society becomes strongly reliant to technologically advanced systems, often located in space such as telecommunication, navigation. Therefore, the conditions and variations in space where these satellites orbit the Earth are important to study and the question arises wether there are influences on such systems or not. We speak of geomagnetic disturbances in this connection. Systems that are susceptible to geomagnetic disturbances are satellites and power grids on Earth. That means that the geomagnetic environment is changed, but as we will see in the later chapters, these disturbances are triggered by our nearest star, the Sun.

It is generally accepted that the term *space weather* refers to the time-variable conditions in the space environment that may effect space-borne or ground based technological systems.

According to the US National Space Weather Programme the definition is: *conditions on the Sun and in the solar wind, magnetosphere, ionosphere and thermosphere that can influence the performance and reliability of space-borne and ground-based technological systems and can endanger human life or health.*

Thus we see that the definitions are slightly different but we want to keep the first one because it also includes other effects apart from the Sun.

[1] see e.g. P. Song, Howard J. Singer, George L. Siscoe, Paul Song, Space Weather, 2001, Am. Geophys. Union

Since we strongly depend on satellite systems and their availability, it is crucial that these systems are in full operation. Moreover, in the worst case, human health or life can also be endangered by space weather. Therefore, there are social and economic aspects of this type of research: one tries to avoid consequences of space weather events by system design or efficient warning and prediction. During the last few years space weather activities have expanded world-wide. Examples for such activities which of course are of national and international interest are:

- US Space Weather Program,
- US-NASA's Living With a Star program,
- ESA's space weather program,
- SWENET, Space Weather European Network,
- SIDC, Solar influences data center at the Royal observatory in Belgium,
- Lund space weather center,
- The Australian IPS Radio and Space Services , the Australian Space Weather Agency,
- The Canadian Space weather program,

and many others (such as the Group in Oulu, Finland). Today, space weather is monitored from a worldwide net of ground stations and from space. Both types of observations are complementary. From space the whole electromagnetic spectrum of the Sun can be observed including UV and X-rays.

An overview about space weather, environment and societies can be found in the monograph by Lilenstein and Bornarel, 2005 [196].

1.2 The Triggers of Space Weather

The main cause for space weather effects is our Sun. It emits light at all wavelengths that reaches the Earth within about 8 minutes as well as a continuous stream of particles which is called the solar wind. During one solar activity cycle which has a period of about 11 years both radiation and solar wind are modulated. The energy of the Sun drives temperature, precipitation, atmospheric circulation, ocean currents, evaporation and cloud cover. The short wavelength radiation (UV, X-rays) triggers many chemical reactions in the upper atmosphere of the Earth and also the ozone level is modulated by solar activity. It was R.C. Carrington who observed on September 1, 1859 a white light flare[2] that erupted from a group of sunspots and in the following night a great aurora was seen down to low geographic latitudes, even from Cuba. On the same night, a great magnetic disturbance was also recorded. For the first time it was recognized correctly, that a change on the Sun might have directly influenced the environment around the Earth.

[2]The observations were reported to the Royal Astronomical Society

The NASA's Sun-Earth Connections program aims to improve our understanding of solar variability and how this transforms into interplanetary space, how e.g. eruptive events on the Sun (like CMEs, Coronal mass ejections) impact geospace, weather and climate. In the long term NASA plans manned missions to the Moon and even Mars and the need of spaceweather forecasts becomes evident.

A software package called CACTUS, Computer Aided CME tracking detects automatically CME events that could be dangerous by scanning through images produced by the SOHO/LASCO satellite [3].

Related to the solar activity are important effects on spacecraft such as spacecraft charging (surface charging and deep discharges) and single event upsets. The effects on humans in space are also to be considered (radiation, particles). Space weather effects also play a rôle on high altitude/high latitude air-flight; cosmic rays penetrate to the lower atmosphere and pose problems to humans and electronic components of modern aeroplanes. Other influences of space weather include radio wave propagation, satellite-ground communications, global satellite-based navigation systems, power transmission systems etc. Changes of the solar irradiance may be one of the causes for climatic changes on the Earth.

Materials located on the exterior of spacecraft in low Earth orbit are subjected to a number of environmental threats, including atomic oxygen, ultraviolet radiation, thermal cycling, and micrometeorid and debris impact. The number of space debris now clearly exceeds the number of meteroids[4].

A compendium of space weather related scientific papers can be found in the book of Scherer, Fichtner and Heber[5]. A book on Solar And Space Weather Radiophysics appeared recently[6].

1.2.1 Examples

Let us give some examples of space weather influences on satellites[7]

- Space Shuttle: numerous micrometeroid/debris impacts have been reported.
- Ulysses: failed during peak of Perseid meteoroid shower.
- Pioneer Venus: Several command memory anomalies related during high-energy cosmic rays.
- GPS: photochemically deposited contamination on solar arrays.

On the Earth we know very well radio fadeouts. The HF communication depends on the reflection of signals in the upper Earth's atmosphere. This layers are strongly influenced by the Sun's shortwave radiation.

[3] see http://sidc.oma.be/products/cactus/index.php

[4] The term meteorite denotes the piece that was fallen to the surface of the Earth

[5] see: Scherer, Fichtner and Heber (Eds.), Space Weather - The Physics behind a Slogan, Lect. Notes in Phys., Springer, 2004

[6] see: Solar And Space Weather Radiophysics: Current Status And Future Developments, D. E. Gray and Ch. U. Keller, ASSL, 2005, Kluwer

[7] see also: The Space Environment, A. C. Tribble, Princeton Univ. Press, 2003

How can we study the propagation of solar disturbances through the interplanetary medium from Earth? A common technique is to measure scintillations in the radio wavelength. Let us consider very distant radio sources like quasars. If the interplanetary medium is not disturbed, the signal from this object is constant in amplitude. But similar to the twinkling of starlight in the visible, the radio signal becomes absorbed and refracted when passing through a plasma cloud emitted from the Sun. By measuring many point sources distributed all over the sky, one gets a map of areas of high scintillation which shows where the plasma wave is propagating.

There are similarities with atmospheric weather, however the most important differences between atmospheric and space weather systems are:

- Meteorological processes are localized; it is possible to make good local weather forecasts. Spaceweather is always global in the planetary scale.
- Space weather events occur over a wide range of time scales: the Earth's magnetosphere responds to solar-originated disturbances within only a few minutes, global reconfiguration occurs within some 10 minutes. Enhanced fluxes of energetic particles in radiation belts decay in time scales of days, months or even longer.
- Spaceweather predictions must rely on the input of just a few isolated measurements of the solar wind and the observations (both ground based and from space) have only a global character sometimes without details.

Therefore, successful space weather activities aiming to make prediction of dangerous events need to be performed on a global scale. Space-borne and ground based observations are complementary.

1.3 Who are the Users of Space Weather?

Presently, the most important users of space weather research are spacecraft engineering, spacecraft operations, RF communications. Spacecraft launchers can make use of exact knowledge of space weather conditions and the re-entry of spacecrafts depends on the atmospheric drag conditions. When the International Space Station, ISS, is in operation, forecasts will become even more important. Other users are telecommunication operators, users of global positioning systems, electric power industry etc. Commercial airlines must be careful with the radiation doses to their crews and passengers.

In 1989 (March 13th) solar activity induced a huge geomagnetic storm causing a saturation in the transformers and the power grid servicing Canada's Quebec province was completely shut down. The blackout resulted in a loss of 19 400 MW in Quebec and 1325 MW of exports. Service restoration took over 9 hours (after R. Thompson, IPS, Radio and Space Service).

Long term variations of space weather are also called space climate . We know that there were periods of reduced solar activity during the past 1000 years (called Spörer Minimum, Maunder Minimum and Dalton Minimum). Though also other

influences such as the eruption of big volcanoes played a role, it is assumed that during these phases the global climate on Earth was cooler than on the average.

Summarizing, the following branches strongly depend on space weather:

- Spacecraft & Aircraft,
- Communication Systems,
- Power Distribution Networks and Pipelines,
- Oil and mineral Prospecting,
- Risks to human health,
- Space weather influence on climate change,
- Insuring against space weather effects.

The effects of space weather on technology infrastructure were discussed in the monograph of Daglis[8].

1.4 Organization of the Book

The book is organized as follows. First we want to give a brief review about the main source of space weather effects, our Sun. The basic physics of the Sun will be discussed since it is essential to understand the mechanisms that cause solar variability. This is necessary in order to make prediction models for space weather forecasts. Then we will speak about the influence of solar variability on the Earth's atmosphere. The last chapters deal about other than solar influences on the conditions in space such as meteoroids, space debris.

The field of space weather and solar physics itself as well as dynamics of space is rapidly evolving. In this second edition new material was included. Additionally, to each chapter recommended textbook references are given. Suggestions from readers of the first edition have been taken into account and are greatly acknowledged.

[8]see: I.A. Daglis, Effects of Space Weather on Technology Infrastructure, 2004, Kluwer

Chapter 2

The Sun a Typical Star

Our Sun is the only star which is close enough to observe details on its surface such as sunspots, faculae, prominences, coronal holes, flares etc., which are all summarized as solar activity phenomena. Therefore, the study of the Sun is important for astrophysics in general. Theories about stellar structure and evolution can be studied in detail on the Sun[1].

On the other hand, the Sun is the driving factor for the climate on the Earth and the structure and shape of the Earth's magnetosphere thus determining and influencing the near Earth space environment. Therefore, the study of solar terrestrial relations is of great importance for our modern telecommunication systems both based on Earth and in space.

2.1 The Sun and Stars

2.1.1 Location of the Sun

More than 99% of the mass of the solar system, to which the Sun, 8 great planets, dwarf planets (such as Pluto) satellites of planets, asteroids, etc. belong, is concentrated in the Sun. The Sun is the nearest star to us and our solar system is located in the Milky Way Galaxy. Our galaxy contains more than 2×10^{11} solar masses (i.e. at least as many stars). The mass of the galaxy can be inferred from the rotation of the system. All stars rotate about the center of the galaxy which is at a distance of about 27 000 light years (Ly) to us [2].

At the location of the Sun in the galaxy, one period of revolution about the galactic center is about 200 Million years. Galaxies in general contain some 10^{11} stars. About 50% of the stars have one or more stellar companions. Up to now more than 150 planetary companions were detected around nearby stars, so called

[1] For textbooks see e.g. Zirin, H., 1988, Astrophysics of the Sun, Cambridge University Press; The Cambridge Encyclopedia of the Sun, K.R. Lang, 2001, Cambridge Univ. Press; A Guide to the Sun, K.H. Phillips, 1995, Cambridge Univ. Press; The Sun, M. Stix, 2002, Springer Verlag

[2] 1 Ly $=10^{13}$ km, the distance light travels within one year propagating through space at a speed of 300 000 km/s

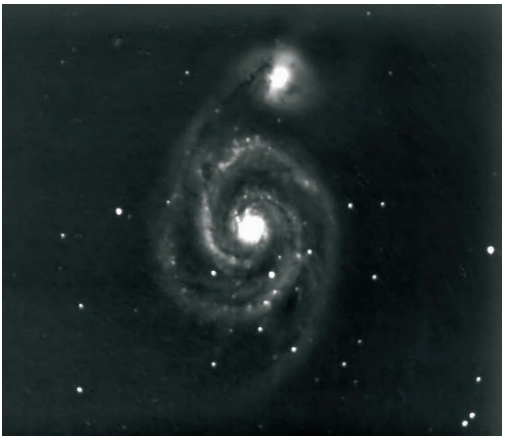


Figure 2.1: A typical spiral galaxy. From a distant galaxy, the Sun would be located in one of the spiral arms. Image: A.H., private observatory.

extrasolar planets. The diameter of our galaxy is about 100 000 Ly. Galaxies are grouped into clusters- our galaxy belongs to the so called local group of galaxies. The small and large Magellanic cloud are two small dwarf galaxies which are satellites of our system. The nearest large galaxy is the Andromeda galaxy which is at a distance of more than 2 Million Ly.

Many galaxies appear as spiral galaxies. Young bright stars are found in the spiral arms, older stars in the center and in the halo of the galaxy. An example is given in Fig.2.1.

2.1.2 Properties of Stars

The only information we can directly obtain from a star is its radiation and position. In order to understand the physics of stellar structure, stellar birth and evolution we have to derive quantities such as stellar radii, stellar masses, composition, rotation, magnetic fields etc. We will just very briefly discuss how these parameters can be derived for stars.

- Stellar distances: a fundamental but not an intrinsic parameter. Stellar distances can be measured by determining their parallax, that is the angle the Earth's orbit would have seen from a star. This defines the astrophysical distance unit *parsec*. A star is at a distance of 1 parsec if the parallax is $1''$. 1 pc = 3.26 Ly.

- Stellar radii: once the apparent diameter of a star is known than its real diameter follows from its distance d. The problem is to measure apparent stellar diameters since they are extremely small. One method is to use interferometers, one other method is to use occultation of stars by the moon

or mutual occultations of stars in eclipsing binary systems. All these methods are described in ordinary textbooks about astronomy.

- Stellar masses: can be determined by using Kepler's third law in the case we observe a binary system. Stellar masses are very critical for stellar evolution, however we know accurate masses only for some 100 stars.

- Once mass and radius are known, the density and the gravitational acceleration follow. These parameters are important for the stellar structure.

- Stellar rotation: For simplicity we can assume that a star consists of two halves, one half approaches to the observer and the spectral lines from that region are blueshifted, the other half moves away and the spectral lines from that area are redshifted. The line profile we observe in a spectrum is a superposition of all these blue- and redshifted profiles and rotation causes a broadening of spectral lines;

- Stellar magnetic fields: as it will be discussed in more detail when considering the Sun, magnetically sensitive spectral lines are split into several components under the presence of strong magnetic fields.

2.1.3 Stellar Spectra, the Hertzsprung-Russell-Diagram

The analysis of stellar radiation is fundamental for the derivation of physical quantities describing a star. Putting a prism or a grating inside or in front of a telescope, we obtain a spectrum of a star. Such a spectrum contains many lines, most of them are dark absorption lines. Each chemical element has a characteristic spectrum.

In the Hertzsprung Russell Diagram (HRD) the temperature of stars is plotted versus brightness. The temperature of a star is related to its color: blue stars are hotter than red stars. In the HRD the hottest stars are on the left side. The temperature increases from right to left. Stellar brightness is given in *magnitudes*. The magnitude scale of stars was chosen such that a difference of 5 magnitudes corresponds to a factor of a 100 in brightness. The smaller the number (which can be even negative) the brighter the star. The brightest planet Venus e.g. has magnitude $-4.^{\mathrm{m}}5$ and the Sun has $-26.^{\mathrm{m}}5$. The faintest stars that are visible to the naked eye have magnitude $+6.^{\mathrm{m}}0$. Since the apparent magnitudes depend on the intrinsic luminosity and the distance of a star absolute magnitudes were invented: the absolute magnitude of a star (designated by $^{\mathrm{M}}$) is the magnitude a star would have at a distance of 10 pc. In the HRD we can plot absolute magnitudes as ordinates instead of luminosities. The relation between m and M is given by:

$$m - M = 5 \log r - 5 \tag{2.1}$$

r is the distance of the object in pc. The Sun has $M = +4.^{\mathrm{M}}5$; seen from a distance of 10 pc it would be among the fainter stars visible with the naked eye.

How can we determine stellar temperatures? Stars can be considered to a very good approximation as *black body* radiators. A black body is a theoretical

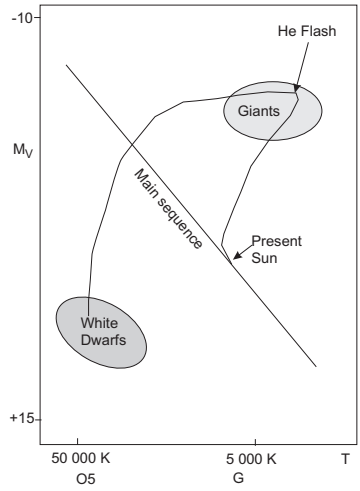


Figure 2.2: Sketch of the Hertzsprung-Russell-diagram with evolutionary path of the Sun.

idealization: an object that absorbs completely all radiation at all wavelengths. The radiation of a black body at a given temperature is given by the *Planck law*:

$$I_\nu = B_\nu = (2h\nu^3/c^2)/\exp(h\nu/kT_\mathrm{S}) - 1 \tag{2.2}$$

thus it depends only on the temperature T_S of the object. Here, I_ν is the intensity of radiation at frequency ν; h, k, c are Planck's constant, Boltzmann's constant and the speed of light. $h = 6.62 \times 10^{-34}\,\mathrm{Js}^{-1}$, $k = 1.38 \times 10^{-23}\,\mathrm{JK}^{-1}$. If that equation is integrated over all frequencies (wavelengths), we obtain a formula for the total power emitted by a black body, Boltzmann law:

$$\int_0^\infty B_\lambda d\lambda = \sigma T^4, \tag{2.3}$$

and for the luminosity of a star:

$$L = 4\pi r^2 \sigma T_\mathrm{eff}^4 \tag{2.4}$$

For the Sun $T_\mathrm{eff} = 5\,785\,\mathrm{K}$. This formula defines the effective temperature of a star. $\sigma = 5.67 \times 10^{-8}\,\mathrm{W/m^2K^4}$ is the Stefan Boltzmann constant.

What is the power emitted per unit area of the Sun's surface? Answer: Put T = 6 000 K we find that the Sun radiates 70 MW per m^2 of its surface[3]

[3]The worldwide nuclear energy generation is about 350 GW. Thus an area of 5000 m^2 on the Sun generates this amount.

Table 2.1: Central wavelength and bandwidth of the UBVRI filter set

Name	Meaning	Central λ	Bandwidth [nm]
U	Ultraviolet	360	66
B	Blue	440	98
V	Visual (green)	550	87
R	Red	700	207
I	Infrared	900	231

By taking the derivative with respect to λ of Planck's Law and setting it equal to zero, one can find the peak wavelength, where the intensity is at maximum:

$$T\lambda_{\max} = 2.9 \times 10^{-3}\,\mathrm{m\,K} \tag{2.5}$$

This is also called *Wien's law.*

At about which wavelength can planets be expected to radiate most of their energy? Answer: Let us assume the temperature of the Earth = 300 K. Then

$$\lambda_{\max} = 2.9 \times 10^{-3}/300 \sim 10\,\mu \tag{2.6}$$

The Sun has a surface temperature of about 6 000 K. At what wavelength does the Sun's spectrum peak? Answer:

$$\lambda_{\max} = 2.9 \times 10^{-3}/6000 \sim 0.5\,\mu = 500\,\mathrm{nm} \tag{2.7}$$

From the spectrum stellar temperatures can be obtained. The temperature derived from the peak wavelength is called Wien Temperature, the temperature derived from the difference of intensity between two wavelengths (=color) Color temperature etc. In order to define color, a filter system must be defined. The most commonly used system is the UBV system which has three bands that are located in the UV (U), blue (B) and visual (V) to measure the intensity I_ν. The luminosity of stars is given in magnitudes which are defined as follows:

$$\text{Magnitude} = \text{const} - 2.5\log(\text{Intensity}) \tag{2.8}$$

The color of a star is measured by comparing its magnitude through one filter (e.g. red) with its magnitude through another (e.g. blue).

E.g. m_V means the magnitude measured with the V filter. Therefore, instead of determining temperatures from the comparison of the spectrum of a star with the Planck law, one can use e.g. color indices. If we calculate B-V, than this value will be (see e.g. table 2.2):

- positive for the cooler star, since it is brighter in V than in B (blue). If the cool star is brighter in V it means that its magnitude has a lower value and therefore B-V is positive.

Table 2.2: B-V colors and effective temperatures of some stars

Star	B-V	Effective T
Sun	+0.6	5 800 K
Vega	0.0	10 000 K
Spica	-0.2	23 000 K
Antares	+1.8	3 400K

- negative for the hotter star. The hotter star is brighter in B than in V, therefore for the magnitudes in these two bands: $m_B < m_V$ and B-V<0.

2.1.4 Stellar Evolution

Stars are not randomly distributed in the HRD:

- Main sequence stars: most stars are found along a diagonal from the upper left (hot) to the lower right (cool).
- Giants, supergiants: they have the same temperature as the corresponding main sequence stars but are much brighter and must have larger diameters (see equation 2.4).
- White dwarfs are faint but very hot objects thus from their location at the lower left in the HRD it follows that they must be very compact (about 1/100 the size of the Sun).

This leads to the question why most of the stars we observe lie on the Main sequence. The answer is quite easy: because this denotes the longest phase in stellar evolution. Let us discuss this briefly for the Sun:

The main steps in the evolution of the Sun are (compare with Fig. 2.2):

- Pre main sequence evolution: from a protostellar gas and dust cloud the Sun was formed and before it reaches the main sequence where it spends most of its life, the contracting Sun has passed a violent youth, the T Tauri phase.
- At the main sequence the Sun changes extremely slowly remaining there about 10^{10} years. In the core H is transformed to He by nuclear fusion.
- The Sun evolves to a red giant, it will expand and the Earth will become part of the solar atmosphere. The expansion starts when all H is transformed to He in the core. Then a H burning shell supplies the energy. The He flash sets in as soon as in the center He burning sets in. The Sun will evolve to a red giant for some 10^8 ys. It will extend beyond the Earth's orbit.
- Finally, the Sun becomes a white dwarf which slowly cools.

Table 2.3: Spectral classification of stars

O	ionized He, ionized metals
B	neutral He, H stronger
A	Balmer lines of H dominate
F	H becomes weaker, neutral and singly ionized metals
G	singly ionized Ca, H weaker, neutral metals
K	neutral metals molecular bands appear
M	TiO, neutral metals
R,N	CN, CH, neutral metals
S	Zirconium oxide, neutral metals

Table 2.4: Effective Temperature as a function of spectral type

Spectral Type	O	B0	A0	F0	G0	K0	M0	M5
$T_{\rm eff}$ [K]	50 000	25 000	11 000	7 600	6 000	5 100	3 600	3 000

During its evolution, the Sun dramatically changes its radius (the subscript $\odot$ denotes the present day value):

1 $R_\odot$ (present Sun) $\rightarrow \sim 10^4\,R_\odot$ (red giant), $\rightarrow 0.01 R_\odot$ (white dwarf).

For space weather long term evolutionary effects are negligible. But it is interesting to investigate them especially for the early Sun (see the chapter on the faint young Sun problem).

For the main sequence stars there exists a relation between their mass and luminosity:

$$L \sim M^{3.5} \tag{2.9}$$

From 2.9 we see that more massive stars are very luminous and therefore they use up their nuclear fuel much more rapidly than low massive stars like our Sun. Massive main sequence stars that are observed today must have been formed in very recent astronomical history[4].

2.1.5 Spectral Classes

According to their spectra, stars can be classified in the following sequence: O-B-A-F-G-K-M. This is a sequence of temperature (see Table 2.4): O stars are hottest, M stars coolest; the number of absorption lines increase from O to M. Some characteristics are given in Table 2.3.

The luminosity of a star depends on a) temperature $\sim T^4$, b) surface which is $\sim R^2$. Since e. g. a K star may be a dwarf main sequence star or a giant, luminosity classes have been introduced. Class I contains the most luminous supergiants, class II the less luminous supergiants. Class III are the normal giants, class IV the sub giants and class V the main sequence.

[4]In some large interstellar nebulae one observes stars that have an age of some 10^5 years

Now we understand the spectral classification of our Sun: it is a G2V star.

2.2 The Sun

2.2.1 Basic Properties

As it has been mentioned, the Sun is a G2V star in the disk of our Galaxy. The mass of the Sun is:

$$M_\odot = 1.99 \times 10^{30}\,\mathrm{kg} \tag{2.10}$$

An application of Kepler's third law gives us the mass of the Sun if its distance is known which again can be derived from Kepler's third law:

$$\frac{a^3}{P^2} = \frac{G}{4\pi^2}(M_1 + M_2) \tag{2.11}$$

In our case a denotes the distance Earth-Sun (150×10^6 km), P the revolution period of the Earth around the Sun (1 year), M_1 the mass of the Earth and M_2 the mass of the Sun. One can make the assumption that $M_1 << M_2$ and therefore $M_1 + M_2 \sim M_2$.

If we know the distance of the Sun and its angular diameter the solar radius is obtained:

$$r_\odot = 6.96 \times 10^8\,\mathrm{m} \tag{2.12}$$

The measurement of the Sun's angular diameter is not trivial; one possibility is to define the angular distance between the inflection points of the intensity profiles at two opposite limbs. Such profiles can be obtained photoelectrically and the apparent semi diameter at mean solar distance is about 960 seconds of arc ($''$). The orbit of the Earth is elliptical and at present, perihelion (smallest distance of the Sun) is in January.

Knowing the mass and radius of the Sun, the mean density can be calculated:

$$\bar{\rho} = 1.4\,\mathrm{g/cm^3} \tag{2.13}$$

The gravitational acceleration is given by:

$$g = GM/R^2 = 274\,\mathrm{m/s^2} \tag{2.14}$$

The *solar constant* is the energy crossing unit area of the Earth's surface perpendicular to the direction from the Earth to the Sun in every second. In SI the units are $\mathrm{W\,m^{-2}}$. UV and IR radiation from the Sun is strongly absorbed by the Earth's atmosphere. Therefore, accurate measurements of the solar constant have to be done with satellites. ACRIM on SMM and ERB on Nimbus 7 showed clearly that the presence of several large sunspots which are cooler than their surroundings depress the solar luminosity by $\sim 0.1\%$. The Variability IRradiance Gravity Oscillation (VIRGO) experiment on the SOHO satellite is observing total solar and spectral irradiances at 402 nm (blue channel), 500 nm (green channel), and

862 nm (red channel) since January 1996 (for a review see e.g. Pap *et al.* (1999) [243]). The solar luminosity is:

$$L_{\odot} = 3.83 \times 10^{26}\,\mathrm{W} \tag{2.15}$$

And the effective temperature:

$$T_{\mathrm{eff}\odot} = 5780\,\mathrm{K} \tag{2.16}$$

2.2.2 Basic Equations

How a ball of gas and plasma, like a star remains stable against gravitational collapse or free expansion? Let us assume a sphere of mass M and radius R. In most cases there are only two forces:

- gravity: acts inward
- pressure: acts outward

Let us consider a shell inside a star, the lower boundary is at r from the center and the upper at $r + \Delta r$. ΔA is a surface element and P_{outer}, P_{inner} denote the pressure at r and $r + \Delta r$. The net force on such a shell is:

$$F_{\mathrm{net}} = F_{\mathrm{grav}} - F_p \tag{2.17}$$

and $F_p = (P_{\mathrm{outer}} - P_{\mathrm{inner}})\Delta A$. From the above equation:

$$F_p = [P(r) + (dP/dr)\Delta r - P(r)]\Delta A = (dP/dr)\Delta r \Delta A \tag{2.18}$$

By dividing the net force F_{net} by $-\Delta m = -\rho(r)\Delta r \Delta A$, we find the equation of motion of the shell:

$$-\mathrm{d}^2 r/\mathrm{d}t^2 = g(r) + [1/\rho(r)](\mathrm{d}P/\mathrm{d}r) \tag{2.19}$$

If the acceleration is set to zero (when there is a balance), then the *hydrostatic equilibrium* becomes:

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{GM(r)\rho(r)}{r^2} \tag{2.20}$$

Therefore, the pressure at depth h must be high enough to support the weight of the fluid per unit area above that depth. Let us derive an estimate for the central pressure of a star. The pressure is given by:

$$P = g\rho h \tag{2.21}$$

At the center $h = R$; from $g = GM/R^2$ we find the central pressure P_c:

$$P_c = \frac{GM\rho}{R} \qquad \rho = \frac{M}{4\pi R^3/3} \tag{2.22}$$

which leads to:

$$P_c = \frac{3}{4\pi}\frac{GM^2}{R^4} \tag{2.23}$$

For the Sun: $M_\odot = 2\times10^{30}$ kg, and $R_\odot = 7\times10^8$ m. This gives $P_c = 3\times10^{14}\,\mathrm{Nm}^{-2}$ compared to the atmospheric pressure at sea level on Earth of $10^5\,\mathrm{Nm}^{-2}$. This is a very crude approximation, since in reality the density increases with depth and the true central pressure of the Sun is 100 times larger than the estimate.

What happens if a star contracts (which will be the case when the hydrostatic equilibrium condition is not established)? According to the *Virial Theorem* half of the gravitational energy which is set free is radiated away and the other half heats the star.

In most phases of stellar evolution, the structure of a star can be determined by the solution of four first order differential equations:

- hydrostatic equilibrium

$$\frac{dP}{dr} = -\frac{GM\rho}{r^2} \tag{2.24}$$

- mass continuity

$$\frac{dM}{dr} = 4\pi r^2\rho \tag{2.25}$$

- gradient of luminosity

$$\frac{dL}{dr} = 4\pi r^2\rho\epsilon \tag{2.26}$$

- temperature gradient

$$\frac{dT}{dr} = -\frac{3\kappa L\rho}{16\pi acr^2T^3} \tag{2.27}$$

In these equations r is the distance from the stellar center, P, ρ, T are the pressure, density and temperature at radius r, M is the mass contained within r, L the energy carried by radiation across r, ϵ the nuclear energy release. The quantities P, ϵ, κ depend on density, temperature and composition. κ is the *opacity* and measures the resistance of the material to energy transport.

2.2.3 Energy Generation in the Sun

In principle, a variety of different energy generating processes can take place in stars. During the formation and contraction of a protostar in an interstellar cloud no nuclear reactions take place and half of the released gravitational energy is radiated away, the other half increases the temperature of the core (Virial Theorem). As soon as the central temperature exceeds about 10^6 K nuclear reactions start. Energy is generated by the fusion of two lighter particles to form a heavier particle whose mass is smaller than the mass of its constituents, the mass defect being transformed into energy according to $E = \Delta Mc^2$.

Let us consider the fusion of H into He. The mass of 4 H is[5]:

[5] 1 AMU= 1/12 of the mass of the Carbon isotope $^{12}\mathrm{C} = 1.66\times10^{-27}\,\mathrm{kg} = 931\mathrm{MeV/c}^2$

Table 2.5: The principal reaction of the pp chain

Reaction Number	Reaction	Neutrino Energy (MeV)
1	$p + p \rightarrow {}^2\mathrm{H} + e^+ + \nu_e$	0.0 to 0.4
2	$p + e^- + p \rightarrow {}^2\mathrm{H} + \nu_e$	1.4
3	${}^2\mathrm{H} + p \rightarrow {}^3\mathrm{He} + \gamma$	
4a	${}^3\mathrm{He} + {}^3\mathrm{He} \rightarrow {}^4\mathrm{He} + 2p$	
4b	${}^3\mathrm{He} + {}^4\mathrm{He} \rightarrow {}^7\mathrm{Be} + \gamma$	
5	$e^- + {}^7\mathrm{Be} \rightarrow {}^7\mathrm{Li} + \nu_e$	0.86, 0.38
6a	${}^7\mathrm{Li} + p \rightarrow {}^4\mathrm{He} + {}^4\mathrm{He}$	
6b	$p + {}^7\mathrm{Be} \rightarrow {}^8\mathrm{B} + \gamma$	
7	${}^8\mathrm{B} \rightarrow {}^8\mathrm{Be} + e^+ + \nu_e$	0...15

$$4 \times 1.008145\,\mathrm{AMU} \tag{2.28}$$

The mass of the resulting He atom is

$$4.00387\,\mathrm{AMU} \tag{2.29}$$

Thus the mass difference ΔM is

$$0.02871\,\mathrm{AMU} \sim 4.768 \times 10^{-29}\,\mathrm{g} \sim 4.288 \times 10^{-12}\,\mathrm{J} \sim 26.72\,\mathrm{MeV} \tag{2.30}$$

and 0,7% of the mass is converted to energy by Einstein's relation[6]

$$E = mc^2 \tag{2.31}$$

If one assumes that the Sun consists of pure hydrogen which is converted into He, then the total energy ($E = 0.007mc^2$) would be 1.27×10^{45} J. The luminosity of the Sun is $L_\odot = 3.8 \times 10^{26}$ J/s thus there would be energy supply for 10^{11} years.

The so called pp chain (Table 2.5) dominates in stars with relatively low central temperatures (between 5 and $15{\times}10^7$ K, like the Sun) and the CN cycle[7] is dominant in stars with higher central temperatures.

The energy production rate, ϵ, for the pp cycle depends highly on the temperature:

$$\epsilon \sim \rho T^5 \tag{2.32}$$

[6] 1 eV=1.6×10^{-19} J

[7] In the CN cycle C acts as a catalyst to convert H into He

2.2.4 Convection Zone

In the solar core nuclear fusion generates the energy which is transported outwards by radiation. At a depth of a third of the solar radius below the solar surface the convection zone starts, where energy is transported outwards by convective motions. This zone occupies only 2% of the solar mass. Hydrogen and He are practically neutral at the solar surface but they are ionized just below the surface. In these ionization zones the ratio of the specific heat at constant pressure (c_p) to the specific heat at constant volume (c_v) is much lower than the value 5/3. This value is appropriate either to a neutral gas or fully ionized gas. Because $c_p/c_v << 5/3$ convection occurs. For the temperature gradient we already have seen that:

$$\frac{dT}{dr} = -\frac{3\kappa L_{\rm rad}\rho}{16\pi acr^2T^3} \tag{2.33}$$

The total luminosity L is:

$$L = L_{\rm rad} + L_{\rm conv} \tag{2.34}$$

Basically, convection can be treated as an instability; if an element of material is displaced upwards, then it continues to rise if it is lighter than its surroundings. By assuming that the rising element moves sufficiently slowly that it is in pressure balance with its surroundings but that at the same time its motion is adiabatic (no heat exchange between the element and the surroundings), then convection occurs if:

$$\frac{P}{\rho}\frac{{\rm d}\rho}{{\rm d}P} < \left(\frac{P}{\rho}\frac{{\rm d}\rho}{{\rm d}P}\right)_{\rm ad} \tag{2.35}$$

If the stellar material is an ideal classical gas with constant ratio of specific heats γ, then:

$$\frac{P}{\rho}\frac{{\rm d}\rho}{{\rm d}P} < \frac{1}{\gamma} \tag{2.36}$$

The theory which is usually used contains a free parameter, the so called *mixing length* l:

$$l = \alpha H_{\rm p} = \alpha \left| P\frac{{\rm d}P}{{\rm d}r}\right| \tag{2.37}$$

where $H_{\rm P}$ is the pressure scale height. It is supposed that α is of order unity. As we will discuss later, information about the depth of the convection zone comes from a detailed study of solar oscillations. Apart from energy transport one has also to consider that in convection zones there is a uniform chemical composition. This prevents any attempt of heavy chemical elements to settle in the Sun's gravitational field.

2.2.5 Model: Internal Structure of the Sun

In this paragraph we give a table showing the variation of temperature, luminosity and fusion rate as a function of increasing distance from the solar center. Such a model can be calculated from the basic set of equations discussed above.

Table 2.6: Solar model: variation of temperature, luminosity and fusion rate throughout the Sun

Radius fraction in $R_\odot$	Radius $[10^9]$ m	Temperature $[10^6]$ K	% Luminosity	Fusion rate [J/kg s]
0	0.00	15.7	0	0.0175
0.09	0.06	13.8	33	0.010
0.12	0.08	12.8	55	.0068
0.14	0.10	11.3	79	.0033
0.19	0.13	10.1	91	.0016
0.22	0.15	9.0	97	0.0007
0.24	0.17	8.1	99	0.0003
0.29	0.20	7.1	100	0.00006
0.46	0.32	3.9	100	0
0.69	0.48	1.73	100	0
0.89	0.62	0.66	100	0

As it will be discussed later, the interior of the Sun can be investigated by the propagation of waves. Solar models computed with mass loss, microscopic diffusion of helium and heavy elements, and with updated physics have been evolved from the pre-main sequence to present day (Morel *et al.*, 1997 [225]); they are compared to the observational constraints including lithium depletion and to the seismic reference model derived by inversion. Microscopic diffusion significantly improves the agreement with the observed solar frequencies and agree with the seismic reference model within $\pm 0.2\%$ for the sound velocity and $\pm 1\%$ for the density, but slightly worsens the neutrino problem. A review on the current state of solar modeling was given by Christensen-Dalsgaard *et al.* (1996) [67] .

2.3 Observing the Sun

2.3.1 General Remarks

In this short chapter we want to give a few examples of modern solar telescopes. Some remarks are also made concerning optical design and features as well as disturbances caused by the Earth's atmosphere.

Earth-based telescopes must contend with image distortion and scintillation caused by atmospheric disturbances as light reaches us from outer space, stars twinkle, images blur and dance when viewed with telescopes or binoculars. This effect worsens as the zenith angle increases. Temperature changes and winds create variations in atmospheric refractive indices resulting in image distortions. This condition is called "seeing", and is a prime consideration in selecting the location of an observatory. Good sites for observatories are the Canary islands or Hawaii located at heights above the inversion. Thus clouds form deeper than the site of the telescope.

Several factors influence the image quality of solar telescopes. Sunlight can heat up the telescope structure and the main optics causing the so called internal seeing. Considering reflecting telescopes, in particular the main mirror absorbs up to 10% of the collected light and its surface may heat up considerably leading to mirror seeing. The most effective measure to prevent internal seeing is to remove the air entirely, the telescope is evacuated. A window at the entrance and exit preserves the vacuum. The main problem here is to have a window with high optical quality which is thick enough to resist air pressure. Helium filling is an alternative to evacuation. The viscosity of He and the dependence of the index of refraction from temperature are lower than for air whereas temperature conductivity is higher. A forced flow of He inside the telescope tube cancels inhomogeneities. The THEMIS telescope (Télescope Héliographique pour l'étude du Magnetisme et des Instabilités Solaires, see Arnaud *et al.* 1998 [12], Mein, 1997 [222]) has a He filled tube. Other possibilities are to construct open telescopes, such as the DOT (Dutch open telescope at La Palma, see Rutten *et al.*, 2000 [266]). For THEMIS a joint between the telescope tube and the dome which has an entrance window of 1 m prevents air exchanges between outside and inside the dome.

To reduce atmospheric turbulence, systems which dynamically control the wavefront deformations effected by the atmosphere are used (adaptive optics)[8]. Under ideal circumstances, the resolution of an optical system is limited by the diffraction of light waves. This so-called *diffraction limit* is generally described by the following angle (in radians) calculated using the light's wavelength and optical system's pupil diameter d:

$$\alpha = \frac{1.22\lambda}{d} \tag{2.38}$$

The turbulent atmosphere blurs images to resolution of 0.5 to 1 arcsec even at the best sites. Adaptive optics (AO) provides a means of compensating for these effects, leading to appreciably sharper images sometimes approaching the theoretical diffraction limit. One technique that has been developed for overcoming atmospheric blurring is speckle interferometry, in which hundreds of very short exposures ("specklegrams") are processed after the observations to reconstruct the unblurred image. However, because the specklegrams must be short exposures and at the same time should have good signal-to-noise, speckle interferometry is limited to imaging very bright objects. All AO systems work by determining the shape of the distorted wavefront, and using an "adaptive" optical element – usually a deformable mirror – to restore the uniform wavefront by applying an opposite cancelling distortion.

2.3.2 Examples of Telescopes

Very briefly some selected examples of optical solar telescopes are given. The Big Bear Solar Observatory (BBSO) (Fig. 2.3) is located at 2000 m elevation in the middle of Big Bear Lake. This site reduces the image distortion which usually occurs when the Sun heats the ground and produces convection in the air just above ground. Turbulent motions in the air near the observatory are also reduced

[8] A review on high spatial resolution solar observations was given by Bonet, 1999 [43]

Figure 2.3: Big Bear Solar Observatory

by the smooth flow of wind across the lake instead of turbulent flow that occurs over mountain peaks and forests. The main instrument is a 65 cm reflector and is currently being upgraded.

In Fig. 2.4 a drawing of a solar vacuum tower telescope is given. Light enters the vacuum tank through a coelostat system and a mirror. The vertical tank is evacuated in order to avoid turbulence in the telescope itself. At NSO, Kitt Peak, the telescope is a 70 cm f/52 system.

The German Vacuum Tower Telescope (VTT) at the Observatorio del Teide, Tenerife, has two coelostat mirrors (80 cm) and the entrance window to the vacuum tank (BK7) has a diameter of 75 cm and a thickness of 7 cm. The primary mirror has 70 cm and the focal length of the system is 45.64 m. The total field of view is 700 arcsec and the scale is 4.52 arcsec/mm.

Other famous solar instruments for observing the Sun in high spatial resolution mode are the Coupole at the Observatoire Pic du Midi, the Swedish La Palma Solar Telescope (SST), THEMIS in Tenerife, the NSO R.B. Dunn Telescope at Sacramento Peak (DST, 30 inch entrance window, 20 inch solar image) and the McMath-Pierce Facility at Kitt Peak..

GREGOR is a cooperation between three German institutes (Kiepenheuer Institut Freiburg, Astrophysikal. Institut Potsdam and Astrophysikal. Institut Göttingen). An existing telescope will be upgraded to a 1.5 m telescope and will permit high resolution solar observations. The 1.5 m mirror is a lightweight mirror consisting of Silicon Carbide (weight only 180 kg) [9].

The SST [10] is the Swedish Vacuum solar telescope located at the Observatorio Roque de los Muchachos in La Palma. It is a 1 m vacuum telescope. A single lens is used thus permitting observations of high image quality through very narrow filters. The Advanced Technology Solar Telescope (ATST) will be a collaboration between 22 institutes[11].

[9] http://gregor.kis.uni-freiburg.de/

[10] //www.solarphysics.kva.se/NatureNov2002/telescope_ eng.html

[11] http://atst.nso.edu

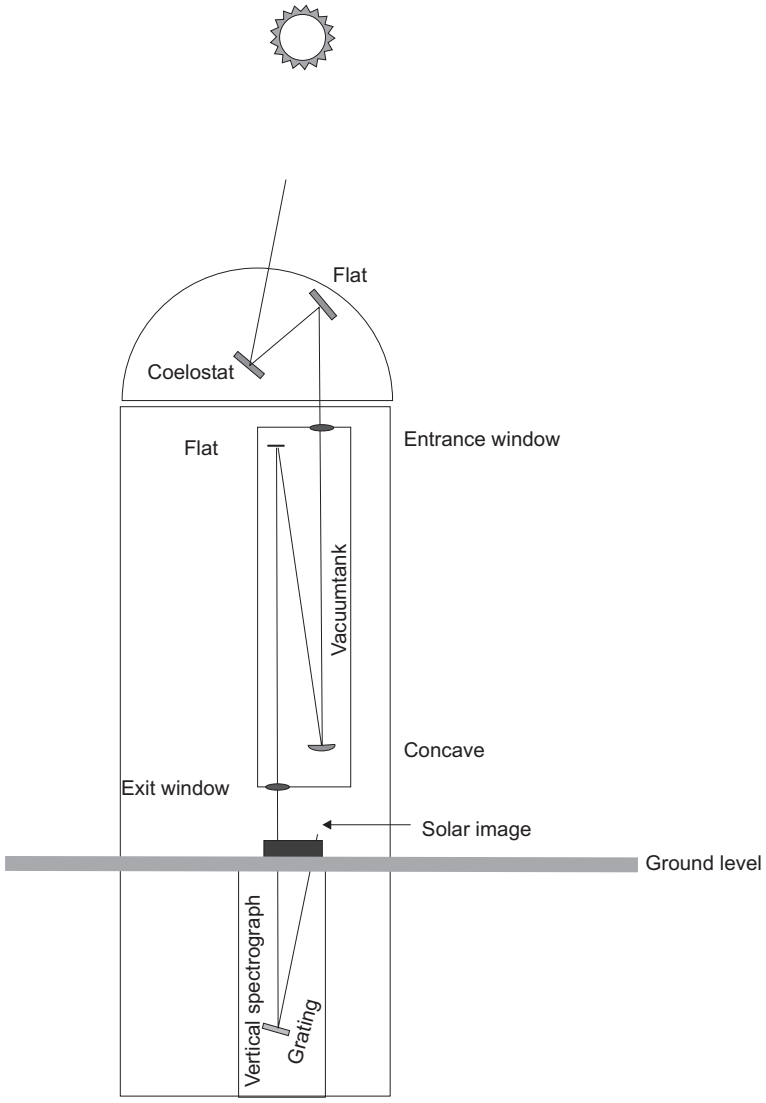


Figure 2.4: Optical path scheme of a vacuum telescope (e.g. Kitt Peak or VTT, Tenerife). Below the ground level, a vertical spectrograph is located. The solar image can be observed at the top of the optical bank that is shown as a black box in the sketch.

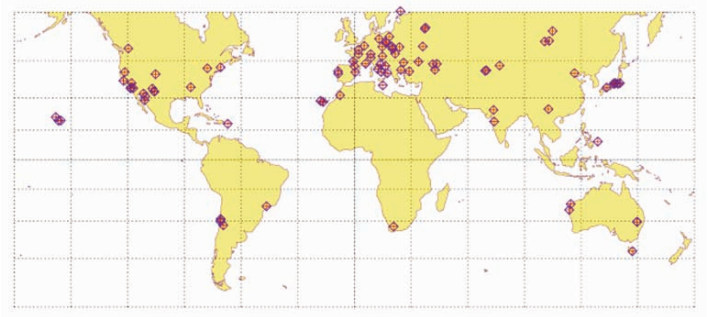


Figure 2.5: Ground-based solar observatories that operate on a routine basis. After K. Reardorn.

Besides these instruments for high resolution studies a number of solar monitoring instruments are distributed over the world (see Fig. 2.5).

A review about solar instrumentation was given by v.d. Lühe (2001) [202]. In the next section we discuss satellite based instruments.

2.3.3 Some Recent Satellite Missions

Here we give three examples of satellites for solar and solar terrestrial research. Further examples will be discussed in the following chapters.

SOHO

The **SO**lar and **H**eliospheric **O**bservatory is a common project being carried out by the European Space Agency (ESA) and the US National Aeronautics and Space Administration (NASA) in the framework of the Solar Terrestrial Science Program (STSP) comprising other missions like CLUSTER and the International Solar Terrestrial Physics Program (ISTP) with Geotail, WIND and Polar. SOHO was launched on December 2, 1995.

SOHO is located at the Lagrangian point L1 about 1.5 Million km away from Earth[12] which permits an uninterrupted view of the Sun. All previous space solar observatories have orbited the Earth, from where their observations were periodically interrupted as our planet 'eclipsed' the Sun.

The scientific objectives of SOHO are the interior of the Sun, the solar atmosphere and the solar wind.

The main instruments are listed in Table 2.7 together with the acronyms.

Further details about SOHO and the instruments can be found e.g. in the review given by Fleck (2001) [98].

A drawing of the spacecraft is given in Fig. 2.6.

[12]This is a stable point in the Sun-Earth system

Table 2.7: Main Instruments on SOHO

Acronym	Name
CDS	Coronal Diagnostic Spectrometer
CELIAS	Charge, Element, and Isotope Analysis System
COSTEP	Comprehensive Suprathermal and Energetic Particle Analyzer
EIT	Extreme ultraviolet Imaging Telescope
ERNE	Energetic and Relativistic Nuclei and Electron
GOLF	Global Oscillations at Low Frequencies
LASCO	Large Angle and Spectrometric Coronagraph Experiment
MDI	Michelson Doppler Imager
SUMER	Solar Ultraviolet Measurements of Emitted Radiation
SWAN	Solar Wind ANisotropies
UVCS	UltraViolet Coronagraph Spectrometer
VIRGO	Variability of solar IRradiance and Gravity Oscillations

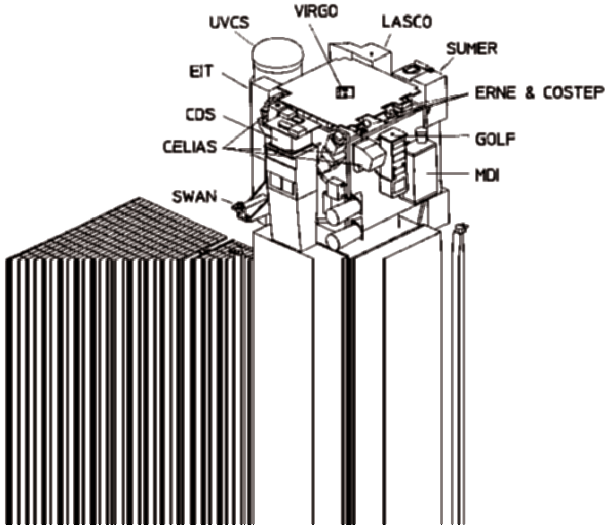


Figure 2.6: Drawing of the SOHO solar observatory (ESA & NASA)

Figure 2.7: Drawing of the RHESSI spacecraft, Artist drawing, NASA

RHESSI

The acronym RHESSI stands for **R**euven Ramaty **H**igh **E**nergy **S**olar **S**pectroscopic **I**mager and it was launched on Feb. 5, 2002. The primary mission objective is to explore the basic physics of particle acceleration and energy release in solar flares. The satellite is located at a circular orbit at 600 km altitude. Observations can be made in the 3 keV to 17 MeV range by using cooled hyperpure germanium crystals. The angular resolution depends on the energy range of the observations: 2 arcseconds to 100 keV 7 arcseconds to 400 keV 36 arcseconds above 1 MeV. During solar flares large amount of energy is released, gas being heated up to several million K radiating X-rays. High energy electrons primarily emit hard X-rays, whereas high energy protons and ions emit primarily gamma rays. RHESSI permits to observe both types of radiation. It seems that the mission will be operational until 2008.

STEREO

This acronym stands for **S**olar **Te**rrestrial **Re**lations **O**bservatory. The mission will consist on two nearly identical space-based observatories - one ahead of Earth in its orbit, the other trailing behind - to provide the first-ever stereoscopic measurements to study the Sun and the nature of its coronal mass ejections, or CMEs. Launch of the STEREO spacecraft is planned for no earlier than September 18, 2006. The main aim will be to study the 3-D structure and extension of coronal mass ejections (CMEs).

2.3.4 Solar Polarimetry

Magnetic fields are the key to understand solar activity phenomena and to predict them. To measure magnetic fields on the Sun, it is inevitable to deal with polarimetry[13].

Electromagnetic radiation consists of oscillations of electric and magnetic fields perpendicular to the direction of propagation. The electric field vector determines the polarization of an electromagnetic wave:

- circular polarization: **E** rotates with its endpoint describing a circle in the plane of polarization, right or left handed, depending on the sense of rotation;
- linear polarization: **E** remains in a fixed position;
- unpolarized: orientation of **E** changes randomly with time.

Polarization can be described mathematically by the four Stokes Parameters: I intensity, Q the linear polarization in the direction of the position angle 0^0, U the linear polarization at 45^0, V the circular polarization. For positive V, the vector **E** is rotates clockwise as seen from the observer.

Usually, the polarization parameters are given relative to the intensity, i.e. Q/I, U/I, and V/I.

The polarization parameters however cannot be measured directly. They have to be measured by using optical devices which are polarization sensitive (linear polarizers, retarders). In a polarimeter, the Stokes vector $S = (I, Q, U, V)$ is transformed into $S' = (I', Q', U', V')$, where the transmitted intensity I' depends on Q, U, V.

Let us consider a simple example: The intensity I of linearly polarized light is measured with a photodetector where the intensity of light generates an electric current proportional to the intensity I. After introducing a linear polarizer into the light beam with arbitrary orientation the measured intensity becomes $I' < I$.

The light is unpolarized if for all orientations of the polarizer the intensity is the same. The orientation of the linear polarizer is varied and thus the intensity of the light measured with the photodetector is modulated.

Because of seeing effects, for precise polarization measurements the modulation frequency must lie above the frequency of the intensity fluctuations caused by disturbances e.g. atmospheric turbulence in solar observations.

The action of linear optical systems on polarized light is described by the Mueller matrices:

$$\mathbf{I}' = \mathbf{MI} \tag{2.39}$$

The intensity of the outcoming beam is a linear combination of all four Stokes parameters of the incoming beam. The Stokes parameters must obey the following conditions:

$$I^2 - Q^2 - U^2 - V^2 \geq 0 \qquad I \geq 0 \tag{2.40}$$

[13] see e.g. K. N. Nagendra, Jan Olof Stenflo, Solar Polarization: Proceedings of an International Workshop Held in Bangalore, India, 12-16 October, 1998, Kluwer, ASSL, 243

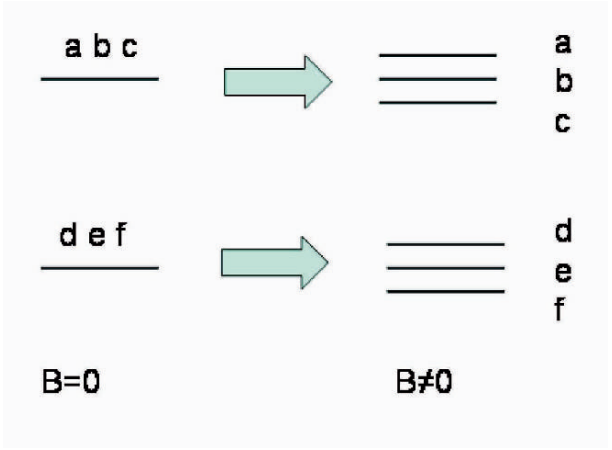


Figure 2.8: Zeeman effect: Without an external magnetic field the energy levels of an electron in an atom a,b,c as well as d,e,f coincide. The presence of a magnetic field splits the energy levels, however not all transitions are possible there are specific transition rules.

When doing solar polarimetry one has to take into account for different effects such as the Earth's atmosphere where time fluctuations produce wavefront distortions, instrumental polarization, spectrograph and detector polarization which all enter as factors in the Mueller matrix.

The two basic effects that can be used for measuring magnetic fields are:

- Zeeman effect: degeneracy of the atomic eigenstates effected by the magnetic field (see Fig. 2.8), splitting of line profiles and characteristic polarization (Stokes V circular, Stokes Q linear). Can be used for magnetic fields $>$ 100 Γ to compete with the microturbulent Doppler broadening of the line profiles. The strength of the Earth's magnetic field is about 0.5 Γ. Also: 1 Γ corresponds to 10^{-4} Tesla .

- Hanle effect: useful diagnostic where the magnetic field is relatively weak (a few to a few tens of Gauss) and where the plasma is sufficiently tenuous that collisional excitation can be neglected in comparison to the radiative excitation of the upper level. It introduces both a rotation of the plane of polarization and a reduction of the net polarization of the scattered light.

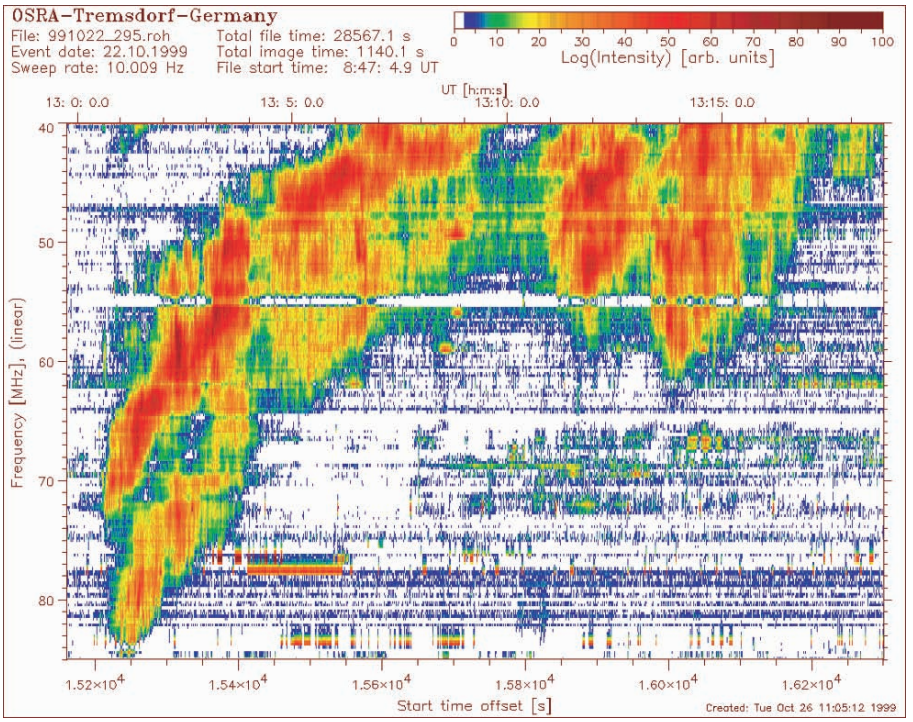


Figure 2.9: Propagation of a wave throughout the outer solar atmosphere. On the abscissa is the time, on the ordinate the frequency. Within 5 min the frequency drifts from 80 MHz to 40 MHz indicating the propagation to the higher corona. Courtesy: H. Aurass, Th. Mann, AIP.

2.3.5 Solar Radio Astronomy

The solar corona is an inhomogeneous, hot, dilute and fully ionized plasma. Its spatial structure is governed by the magnetic field. Plasma processes associated with solar activity take place on small temporal and spatial scales and reveal themselves by electron acceleration (up to a few MeV) which can emit radio radiation. These are called non thermal electrons. Radio radiation from the Sun was first detected in 1942.

From a simple derivation from electrodynamics it follows, that there exists a plasma frequency, ω_P, and that propagating waves through such a plasma are only possible for $\omega_{\text{waves}} < \omega_P$

$$\omega_P = \sqrt{\frac{4\pi n e^2}{m_e}} \tag{2.41}$$

where n is the number of electrons per unit volume and ω_P the *plasma frequency*. Let us consider the propagation of an electromagnetic wave throughout the Sun's outer layers. Here, the density decreases from the chromosphere to the corona.

Therefore, electromagnetic waves with higher frequencies originate in deeper layers and by observing the Sun in different frequency channels one can measure the propagation of a plasma wave through the atmosphere (see Fig. 2.9).

Possible emission mechanisms at cm- and mm-wavelengths are:

- The quiet Sun component of the radio emission is from thermal emission from the hot ionized gas.
- At a frequency of 100 GHz (0.3 cm), the emission originates at the same height in the photosphere than at visible wavelengths. At 1.4 GHz (21 cm) the emission originates from the top of the chromosphere (corresponding to a black body at 100 000 K) and at longer wavelengths (e.g. 300 cm, corresponding to 0.1 GHz) the emission arises from the corona (1-2 Million K blackbody). Thus, the size of the Sun varies when measured at different wavelengths.
- The slowly varying component also has thermal origin and arises from regions above sunspots where the electron density is higher.

The **Lo**w **F**requency **A**rray for **R**adio Astronomy (LOFAR) is an interferometric array of radiotelescopes located in Germany and the Netherlands permitting observations in the 10 - 240 MHz range. The **A**tacama **L**arge **M**illimeter **A**rray (ALMA) is an international collaboration between Europe and the North America to build a synthesis radio telescope that will operate at millimeter and submillimeter wavelengths and it will be able to study gas dynamics around protostars. Moreover, ALMA will be the largest solar observing facility. The instrument is much more sensitive in the mm domain than RHESSI in X-rays. This will permit to study the process of acceleration in flares. In general, the mm and submillimeter emission of the quiet Sun is thermal emission from the chromosphere. Furthermore, studies between 100 and 400 km height above the solar surface can be made and therefore the propagation of solar oscillations can be investigated.

A general introduction to solar radio astronomy can be found in the book of Krüger[14].

2.4 Neutrinos-Testing the Solar Interior

Neutrinos are produced during several phases of the pp reaction[15]. Therefore, they provide a diagnostic tool for the occurrence of the reactions and can serve as a test whether theoretical models of the solar interior are in accordance with observations.

2.4.1 General Properties

The β decay lead to the detection of neutrinos. An example is a β^- decay where a neutron in an atom's nucleus turns into a proton, p, electron e^- and an antineu-

[14] A. Krüger, Introduction to solar radio astronomy and radio physics (Geophysics and astrophysics monographs), D. Reidel, 1979

[15] for a general description see also: Haxton (2001) [129]

trino ν_e like the decay of tritium or ^{3}H (halflife 12.33 years) into ^{3}He:

$$^3\mathrm{H} \rightarrow ^3\mathrm{He} + \mathrm{e}^- + \nu_\mathrm{e} \tag{2.42}$$

During β^+ decay, a proton in an atom's nucleus turns into a neutron, a positron and a neutrino.

The neutrino was first invented as an ad hoc hypothesis, in order to save the laws of conservation of energy and momentum from falsification around 1930. Wolfgang Pauli proposed to explain this discrepancy by postulating that an additional, invisible particle was emitted along with the electron, carrying away the missing energy and momentum. This "ghost particle" was named *neutrino*.

There exist three families of elementary particles, each family consisting of two quarks, and two leptons.

Quarks are constituents of protons and neutrons. Lepton is the collective term for electrons and neutrinos and their relatives in the other families. The electron and the electron-neutrino ν_e make up the lepton pair of the first family. In the other two families, the electron-equivalents are called muon μ and tau τ, each with their neutrino partner, called mu-neutrino ν_μ and tau-neutrino ν_τ. The ν_τ was discovered in 1988.

2.4.2 Solar Neutrinos

As we have seen, neutrinos are produced in the first reaction of the pp chain having an energy between zero and 0.42 MeV. The maximum energy for the neutrinos from the decay of ^{8}B is about 15 MeV. All neutrinos interact very weakly with matter, the probability of absorption increases with their energy. The rare ^{8}B neutrinos are more likely to be absorbed. The absorption cross section is the effective area offered by a target particle to a beam of incident particles. For neutrinos the cross section σ is:

$$\sigma \sim 10^{-50}\,\mathrm{m}^2 \tag{2.43}$$

When we compare this value to the cross sections in atomic and nuclear physics which are about $10^{-20}...10^{-30}\,\mathrm{m}^2$ we see that neutrinos can penetrate the whole Sun without being absorbed and therefore they can be used to test our models. The distance between collisions, the mean free path l, if the target particles have a number density n m^{-3}, is given by:

$$l = \frac{1}{n\sigma} \tag{2.44}$$

For a solid target one has $n \sim 10^{29}$ and therefore $l \sim 10^{21}$ m for neutrinos. So neutrinos have an extremely large mean free path. How can we detect them? There are many neutrinos coming from the Sun passing the Earth: about $10^{15}\,\mathrm{m}^{-2}\mathrm{s}^{-1}$. With the cross section and the number density given above, the number of detections N would be:

$$N \sim 10^{-6}\,\mathrm{m}^{-3}\,\mathrm{s}^{-1} \tag{2.45}$$

That means about one neutrino per month per cubic meter of the detector.

2.4.3 Solar Neutrino Detectors

The first experiment to detect solar neutrinos was a ^{37}Cl reaction with neutrinos resulting in ^{37}Ar which is unstable and decays to ^{37}Cl. The decay of Ar can be measured.

$$\begin{aligned} {}^{37}\mathrm{Cl} + \nu_{\mathrm{e}} &\rightarrow {}^{37}\mathrm{Ar} + e^{-} \\ {}^{37}\mathrm{Ar} &\rightarrow {}^{37}\mathrm{Cl} + e^{+} + \nu_{e} \end{aligned} \tag{2.46}$$

Only neutrinos with energies > 0.8 MeV can be detected by this reaction. This rules out the most numerous low energy neutrinos (first reaction in the pp chain). The ^{37}Cl is in a tank containing 10^5 gallons of C_2Cl_4 perchlorethylene in the Homestake Gold Mine in Lead, South Dakota. The experiment has to be placed deep below the surface to avoid contaminating reactions produced by cosmic rays. Ar is an inert gas, one can extract it from the tank and observe its decay elsewhere.

Neutrino detections are measured by the solar neutrino flux unit defined by:

$$1\,\mathrm{SNU} = 10^{-36}\,\mathrm{interactions\,s^{-1}\,target\,atom^{-1}} \tag{2.47}$$

Since the experiment contains about 2^{30} ^{37}Cl atoms one has to expect one detection every 5×10^5 s.

Theoretical models of the Sun predict the following count rates:

Bahcall, Pinsonneault (1992) [22] 8.0 ± 3.0 SNU,

Turck-Chièze, Lopes (1993) [319] 6.4 ± 1.4 SNU.

However, the measured flux is:

$$2.28 \pm 0.23\,\mathrm{SNU} \tag{2.48}$$

As it has been stressed already, the chlorine experiment is (according to standard solar model predictions) sensitive primarily to neutrinos from the rare fusion reaction that involves ^{8}B neutrinos which are produced in only 2 of 10^4 terminations of the pp cycle. In a conference held in Brookhaven, 1978, it was therefore suggested to design new experiments that are sensitive to the low energy neutrinos from the fundamental pp reaction.

In the Ga experiment, neutrinos with an energy $\geq 0.2332\,\mathrm{MeV}$ can initiate the reaction:

$$\nu_{\mathrm{e}} + {}^{71}\mathrm{Ga} \rightarrow \mathrm{e}^{-} + {}^{71}\mathrm{Ge} \tag{2.49}$$

Therefore, many of the pp neutrinos are included. The SAGE is a Russian/American experiment and uses 60 tons of metallic Gallium. The GALLEX experiment is a European experiment located underground in Italy. It uses 30 tons of Ga in a $GaCl_3HCl$ solution. More than half of the neutrinos that can be detected with this experiment come from the second most important contribution of the pp chain, from ^{7}Be. Again the results are inconsistent with theoretical predictions with a discrepancy by a factor of about 2. However, they provided a first experimental indication of the presence of pp neutrinos.

The Kamiokande experiment uses a large tank of pure water sited underground and its aim was to study the possible decay of the proton. The half life of a p

is $\sim 10^{30}$ yr. The neutrino detector picked up a number of neutrinos from the explosion of the supernova SN 1987A in the Large Magellanic Cloud, which is a neighbor of our galaxy. In an updated version (Kamiokande II) 0.68 kilotons of water were used and neutrinos from ^{8}B above 7.5 MeV can be detected. The water experiment Kamiokande detects higher energy neutrinos (above 7 MeV) by neutrino-electron scattering ($\nu + e \rightarrow \nu' + e'$) and according to the standard solar model the ^{8}B decay is the only important source of these higher-energy neutrinos. The experiment clearly showed that the observed neutrinos come from the sun because the electrons that are scattered by the incoming neutrinos recoil predominantly on the direction of the sun-earth vector. The relativistic electrons are observed by the Cerenkov radiation they produce in the water detector.

The results of the gallium experiments, GALLEX and SAGE gave an average observed rate of 70.5 ± 7 SNU. This is in agreement with the standard model by the theoretical rate of 73 SNU that is calculated from the basic pp and pep neutrinos. The ^{8}B neutrinos which are observed above 7.5 MeV in the Kamiokande experiment, must also contribute to the gallium event rate. This contributes another 7 SNU, unless something happens to the lower energy neutrinos after they are created in the Sun. Thus the Ga experiments are in accordance with predictions if we exclude everything but the pp neutrinos. This is sometimes called the third neutrino problem.

The calculated pp neutrino flux is approximately independent of solar models; it is closely related to the total luminosity of the sun.

Summarizing the the neutrino problem we can state:

- smaller than predicted absolute event rates in the chlorine and Kamiokande experiments.
- incompatibility of the chlorine and Kamiokande experiments,
- very low rate in the Ga experiment which implies the absence of ^{7}Be neutrinos although ^{8}B neutrinos are present.

Solar neutrino experiments are currently being carried out in Japan (Super Kamiokande,Takata (1993) [309], Totsuka (1996) [313]), Canada (SNO, Sudbury, using 1 kiloton of heavy water; Hargrove and Paterson, 1991 [127]) and in Italy (BOREXINO, ICARUS, GNO (Gallium Neutrino Observatorium), each sensitive to a different energy all working in Gran Sasso, Arpesella *et al.* (1992) [13]), in Russia (SAGE, Caucasus) and in the United States (Homestake). The SAGE, chlorine and GNO work radiochemical, the others electronic (recoil electrons produced by the neutrino interactions using Cherenkov effect).

The Sudbury Neutrino Observatory (SNO) (see also Krastev (2002) [173]) is located at 6800 feet under ground in a mine in Sudbury, Ontario. The neutrinos react with heavy water producing flashes of light (Cerenkov radiation). The principle is as follows:

$$\nu_e + d \rightarrow p + p + e^- \tag{2.50}$$

As the neutrino approaches the deuterium nucleus d a heavy charged particle of the weak force (called the W boson) is exchanged. This changes the neutron in

deuterium to a proton, and the neutrino to an electron. The electron, according to mechanics, will get most of the neutrino energy since it has the smaller mass. Due to the large energy of the incident neutrinos, the electron will be so energetic that it will be ejected at light speed, which is actually faster than the speed of light in water. This causes the optical equivalent of a "sonic boom", where a "shock wave of light" is emitted as the electron slows down. This light flash, called Cherenkov radiation, is detected.

The current status of solar neutrino experiments was reviewed by Suzuki (1998) [303] and by Shibahashi (2003) [282].

2.4.4 Testing the Standard Solar Model

We will speak about helioseismology in a later chapter, however in this context it should be noted that results from helioseismology have increased the disagreement between observations and the predictions of solar models with standard neutrinos. Helioseismological measurements demonstrate that the sound speeds predicted by standard solar models agree with high precision with the sound speeds of the sun inferred from measurements. This leads to the conclusion that standard solar models cannot be wrong to explain the discrepancy.

The square of the sound speed is:

$$c^2 \approx T/\mu \tag{2.51}$$

where T is the temperature and μ the mean molecular weight. Sound speeds can be determined with the aid of helioseismology to a very high accuracy (better than 0.2% rms throughout nearly the whole sun). Thus one can estimate tiny errors in the model values of T and μ as measurable discrepancies in the precisely determined helioseismological sound speed:

$$\frac{\delta c}{c} \simeq \frac{1}{2}\left(\frac{\delta T}{T} - \frac{\delta \mu}{\mu}\right) \tag{2.52}$$

The quantitative agreement between standard model predictions and helioseismological observations rules out solar models with temperature or mean molecular weight profiles that differ significantly from the standard values. This observational agreement rules out in particular solar models in which deep mixing has occurred. The best agreement is obtained when including the effect of particle diffusion-selective sinking of heavier species in the sun's gravitational field. Models without taking into account of this effect have rms discrepancies between predicted and measured sound speeds as large as 1% (e.g. Turck-Chièze and Lopez (1993) [319] whereas models including this effect have rms discrepancies of 0.1% (Bahcall *et al.*, 1997 [23]).

The sound-speed profile in the Sun was determined by carrying out an asymptotic inversion of the helioseismic data from the Low-Degree (l) Oscillation Experiment (LOWL), the Global Oscillation Network Group (GONG), VIRGO on SOHO, the High-l Helioseismometer (HLH), and observations made at the South Pole (Takata and Shibahashi, 1998 [310]). Then the density, pressure, temperature, and elemental composition profiles in the solar radiative interior were

deduced by solving the basic equations governing the stellar structure, with the imposition of the determined sound-speed profile and with a constraint on the depth of the convection zone obtained from helioseismic analysis and the ratio of the metal abundance to the hydrogen abundance at the photosphere. Using the resulting seismic model, neutrino fluxes were estimated and the neutrino capture rates for the chlorine, gallium, and water Cerenkov experiments. The estimated capture rates are still significantly larger than the observation.

Solar models with helioseismic constraints and the solar neutrino problem are discussed in Watanabe and Shibahashi (2001) [331] and Roxburgh (1998) [263].

Is there a correlation between neutrino fluxes and solar activity? On the basis of an analysis of the ^{37}Ar production rate at the Homestake station for the period 1970-1990, Basu (1992) [28] found that the solar neutrino flux varies with time in proportion to the solar wind flux. However, Walther (1999) [330] found that there exists no significant correlation between the Homestake neutrino data up to run 133 and the monthly sunspot number, according to a test that is based on certain optimality properties for this type of problem. It is argued that the reported highly significant results for segments of the data are due to a statistical fallacy.

2.4.5 Solution of the Neutrino Problem

Neutrino Oscillations

How to explain this discrepancy between observations and theory? One explanation comes from particle theory itself. There are three conserved quantities, called electron, muon and tauon lepton numbers and correspondingly three types of neutrinos. There are however indications that some modifications to this standard model are required. These involve that the neutrinos have small masses and that the neutrinos can transform from one type to another. The Mikheyev-Smirnov-Wolfenstein effect (MSW) explains such neutrino oscillations and by the above mentioned experiments we can only detect electron neutrinos. Another explanation of the discrepancy is that the flux is variable during the solar cycle. This might be explained if the neutrinos possess a magnetic moment and if interaction with the solar magnetic field is possible.

Let us give a very simplified explanation of neutrino oscillations. An indispensable, but counterintuitive, concept in quantum mechanics is that of superposition. Suppose a certain particle has a property that can have several different values; the classic example is that of Schrödinger's cat, but let us consider a more practical one: ordinary playing cards have the property 'suit', with the four possible values 'spades', 'hearts', 'diamonds', and 'clubs'. In ordinary non-quantum life, each individual card has a well-defined *suit*. However a quantum card may be in a mixed state, a superposition of, say 30% spades, 60% hearts, and 10% clubs. When you check which suit that card belongs to, you have a 30% chance of finding that it's a spade, 60% chance of finding it's a heart, and so on. Note that this is not just a matter of your ignorance of the card's "true" suit – the point is, it doesn't have a single well-defined suit until you check it.

In particle physics, the equivalent of the suits are the three families, discussed above. A neutrino may belong to any one of the three families, making it ν_e, or a ν_μ, or a ν_τ. Or, it may be a superposition of the three family flavors, mixed in some proportions. Now, the standard model assumes that the neutrinos emitted from the sun are in a pure ν_e state, without mixing. This can be understood with the quantum mechanical concept of eigenstates. This is well known for the K meson. An eigenstate is a state that is recognized as pure, non-mixed, without superposition, in a certain context. In quantum mechanics different interactions recognize and interact each with a different set of eigenstates for the particles. Try to apply this to a card game. In different games a heart would become a spade etc. For most particles and interactions the different eigenstates are identical. This is not the case for the weak interaction. The weak eigenstates of quarks are different from their strong/electromagnetic eigenstates. The K0 mesons are produced in strong interactions of quarks, but decay through weak interactions of their constituent quarks. Thus, the production eigenstates are different from the travel/decay eigenstates of the K0. As far as the weak interaction is concerned, leptons are expected to behave in the same manner as quarks. If neutrinos do have a tiny mass, and different neutrinos have different masses, they will behave in the same way as K0 mesons. They will be produced in a weak-interaction eigenstate, but travel in a mass eigenstate. The mass eigenstate may be different from the weak eigenstate. The weak-interaction eigenstates are the three neutrino flavors discussed earlier: ν_e, ν_μ, ν_τ. When they arrive and interact with our detectors, they do not arrive as the original weak eigenstate in which they were produced, but as a mixture of two or more flavors. This is a potential solution to the solar neutrino problem, since the experiments measure an apparent disappearance of electron-neutrinos, without measuring the other flavors. If the neutrinos oscillate from the 100% ν_e that they are produced as in the sun, to a mixture with around 40% ν_e electron-neutrino and 60% some other neutrinos, we get an agreement with experimental data.

Neutrino oscillations and the solar neutrino problem are discussed by Haxton (2001) [129] .

The search for neutrino decays during the 1999 solar eclipse is discussed in Cecchini *et al.* (2000) [60] involving the emitted visible photons, while neutrinos travel from the Moon to the Earth.

Alternate Solar Models

Other suggestions to solve the neutrino problem are:

1. There is an additional force resisting gravity in the solar interior which reduces the central pressure and temperature- maybe rapid rotation, strong internal magnetic fields. Since the pp chain is strongly dependent on temperature, this might explain a different SNU.

2. The Sun contains a central black hole or neutron star. There occurs a gravitational release from accretion providing much of the radiated energy.

3. The surface chemical composition is not typical of the interior composition.

4. Waves or weak interacting particles contribute to the energy transport.

Today the most likely solution of the neutrino problem are the neutrino oscillations suggesting that our solar model is quite correct. Bahcall and Davis (2000) [21] gave a recent review about the solar neutrino problem and suggest further experiments.

2.5 Helioseismology-Solar Oscillations

2.5.1 Observations of Oscillations

In the 1960s the five minutes oscillations were detected on the solar surface. These are vertical oscillations with a strongly varying amplitude but a period of five minutes, maximum velocities about 0.5 km/s towards or away from the observer. The pattern persists for about half an hour (six cycles of wave motion), then fades away but a similar pattern emerges. It was realized that these oscillations could be understood in terms of a superposition of many normal modes of solar oscillations.

Let us consider one analogy: For seismic waves on the Earth one usually has only one source of agitation - an earthquake. For the Sun, there are many sources of agitation of solar "seismic" waves; these sources of agitation causing the solar waves are processes in the larger convective zone. Because there is no single source, we can treat the sources as a continuum, so the ringing Sun is like a bell struck continuously with many tiny sand grains.

How can we measure solar oscillations? Let us briefly describe the main principles of a Dopplerimager. Consider the intensity profile of an absorption line. If the material from which this absorption line is emitted moves away from the observer, the line will be redshifted according to the Doppler effect. We can use this effect to produce velocity images of the solar surface. The light from the Sun is sent through a filter that alternates between letting through light from a narrow range of wavelengths on either side of the center of the line. The two light intensities are measured at every point on the solar surface using an imaging camera. The difference between the two intensities changes when the spectral line shifts, and therefore that difference is a measure of the velocity (see Fig. 2.10).

Waves travelling from the interior of the Sun up to the surface are reflected back again at the surface boundary. Imagine a wave normal to the surface of the Sun and travelling towards the center. As the wave travels deeper into the interior, the temperature increases and the wave is progressively refracted away from the normal until it turns around and returns to the surface. At the Sun's surface the sharp density gradient causes subsequent reflection and the wave heads back into the Sun. Thus the Sun is a *resonant cavity*, standing waves are created. The more often a wave returns to the surface, the less deeply it penetrates before being turned back and conversely, waves reflected only a few times from the surface probe much deeper into the Sun.

The Global Oscillation Network Group (GONG) project is a community-based program to conduct a detailed study of solar internal structure and dynamics using helioseismology, by means of observations from a network of six stations (Learmont (Australia), Mauna Loa (Hawaii), Big Bear (US, California), Cerro Tololo (Chile), Teide Observatory (Tenerife, Spain), and Udaipur (India)) spread around the World in order to void gaps due to sunset. To measure solar oscillations

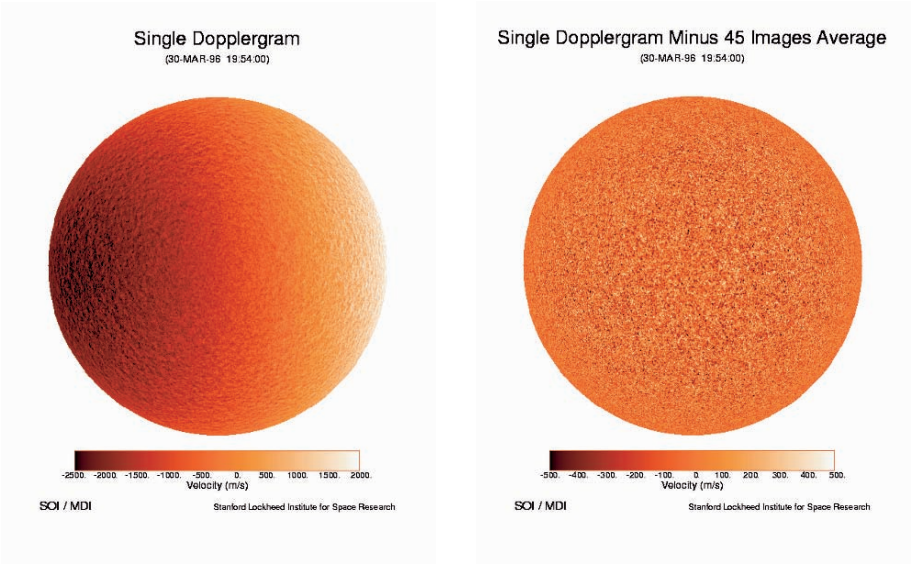


Figure 2.10: SOHO-MDI Dopplerimage; left: the rotation of the Sun is clearly seen; right: the rotation of the Sun was eliminated and therefore only velocities due to granulation and supergranulation are seen.

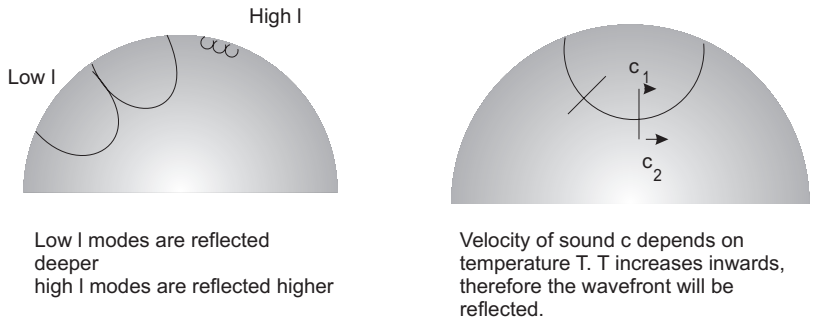


Figure 2.11: Left: waves with low and high l; the low l modes are reflected deeper than the high l modes. Right: Explanation how the waves are reflected in the solar interior. The wavefront (normal to the propagation) is deflected since the sound velocity is higher in deeper layers ($c_2 > c_1$).

one takes a sequence of images of the oscillation pattern at fixed time intervals. The shorter these time intervals between the images, the easier it is to identify the oscillations. A big problem in such a project is the enormous amount of data. Each station in the network produces more than 200 megabytes of data every day. Details about the instrument used (a Fourier Tachometer) can be found in Beckers *et al.* (1978) [30].

The BiSON (Birmingham Solar Oscillations Network) project also has six observatories, most of which are automated. As it is explained above, the GONG observatories measure the motions on the solar surface caused by the oscillations. The BiSON observatories do so as well, but unlike the GONG network they measure an average velocity over the solar surface (the Sun is observed as a point source, if it were a star). The measurements therefore are sensitive only to oscillation patterns with very big wavelengths: all smaller-scale patterns are suppressed by being averaged. The two techniques for GONG and BiSON are therefore complementary.

2.5.2 Modes of Oscillations

There are two different types of oscillations depending on the restoring force.

- p- modes: the restoring force is the pressure;
- g-modes: the restoring force is the gravity.

There exist also surface waves which are called f-modes. The p-modes have frequencies between 1 hour and two minutes and include the five minutes oscillations discussed above. The g-modes have much longer periods than the p-modes. It can be shown that they are trapped in the solar interior beneath the convection zone. The energy generated in the sun is first transported by radiation and then at a depth of about 200 000 km by convection. In this convection zone the amplitudes of the g-modes are damped exponentially and thus it is extremely difficult to observe them at the solar surface. Amplitudes would be expected of a few cm/s to mm/s but the frequency of these modes would contain valuable information about the solar core (Turck-Chièze *et al.*, 2004 [318]. Duvall, 2004 [83], suggested a new method (time distance helioseismology) to detect g modes.

How can we describe the solar oscillations? First we must make some simplifications. We assume that the sun is strictly spherical. This will provide a spectrum of oscillation frequencies which will be modified by a) rotation and b) magnetic fields. A second approximation is that the oscillations are adiabatic. This approximation is valid since the oscillation period is in general much smaller than the relevant thermal timescale. A third approximation is that we neglect a change of the gravitational field of the Sun during the oscillation. This is not true for radial oscillations: in radial oscillations all matter at any solar radius moves inward or outward in phase. However if we consider nonspherical modes at short wavelengths in the horizontal direction this is again a good approximation.

Any oscillation can be described by introducing three quantum numbers n, l, m. The meaning of these numbers is as follows:

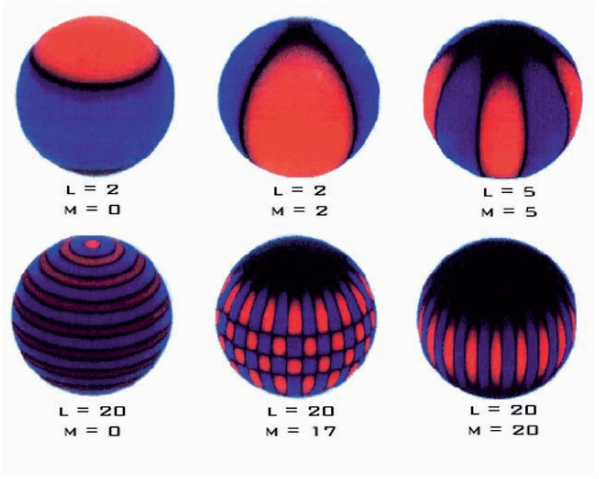


Figure 2.12: Examples of several modes

- n denotes the number of points in the radial direction at which the amplitude of the oscillation vanishes.
- l, m determine the angular behavior of the oscillation over the surface of the Sun. In addition we have the relation $-l \leq m \leq +l$.

If P_l^m denotes the associated Legendre function which can be given in an analytical form, the inward or outward motion of points on the surface is related to the value of the real part of the function

$$P_l^m(\cos\Theta)\exp(im\phi) \tag{2.53}$$

where Θ, ϕ are spherical polar coordinates. If l, m are low, there is a relatively small number of patches on the solar surface (which oscillate with different directions of radial velocity). If l, m are large, there is a very large number of such patches. We speak of a high degree model if l is large and conversely if l is small. Most of the observable p-modes have periods between 2 and 10 minutes with 5 minutes as a characteristic value. These p-modes are trapped near to the solar surface and in the solar interior. For high values of l the modes are trapped close to the surface. In general the oscillation frequency of any mode depends on the internal properties of the Sun in the region which the mode can propagate.

The $l-\nu$ diagram (Fig. 2.13) is fundamental for helioseismology. This diagram shows how much acoustic energy there is at each frequency for every one of the spatial modes of oscillation. A musical instrument should be tuned to a single frequency and a few harmonious overtones, the Sun resonates in tens of millions of ways all at the same time. The frequency ν of each mode reveals a slightly different part of the Sun's interior. The spatial modes are identified from patterns on the Dopplergrams that are made each minute. The frequencies are very low compared

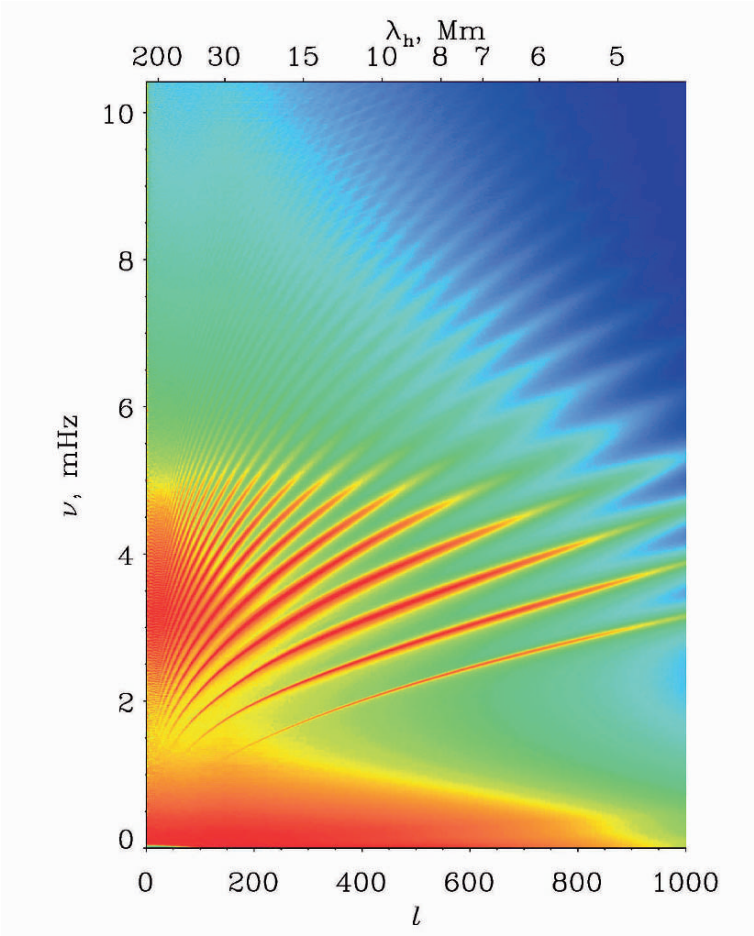


Figure 2.13: l-ν diagram from MDI high-cadence full disk data shows mode frequencies up to 10 mHz and l=1000.

to sound waves we are used to hearing. Most of the power is concentrated in a band near 3 mHz, that's one oscillation every 5 minutes [16]. Higher frequencies aren't trapped inside the Sun, so they don't resonate. Modes with lower ν disappear in the background noise. The spatial scale of the modes is indicated by the angular degree l telling how many node lines there are in the pattern at the surface of the Sun. The l=0 modes are 'breathing' modes where the whole surface of the Sun moves in and out at the same time. Higher order modes divide the surface into a pattern like a checker board, where adjacent squares move in different directions at any given time. A mode of a particular degree, l, at the surface can be associated with resonances having any number of nodes in the radial direction inside the Sun. The number of radial nodes is called the order. The curved lines in the figure are associated with different radial orders. For a given order (line) the

[16]Sound waves we can hear vibrate from tens to thousands of times per second

frequency decreases with increasing spatial degree. For a give degree, the frequency increases with order. In the Fig. 2.13, the lower left corner is most closely related to what is happening in the core of the Sun. Moving up in frequency or degree tells more about what is happening near the surface. Because sound waves of a particular degree can travel in different directions the lines appear relatively broad. If the material through which any of these modes is travelling is moving, then the measured frequency of the mode is affected. The rotation of the Sun causes the biggest frequency shift and makes the lines shown in the figure broad (frequency shifting). Other motions within the Sun along the path taken by the waves cause different types of frequency changes. Analysis of these frequency changes reveals the internal motions of the Sun.

2.5.3 Theory of Solar Oscillations

Let us briefly describe the basic theory of solar oscillations. For an overview of this rapidly evolving topic see also Christensen-Dalsgaard, 2004 [66]. We use the basic equations:

$$\rho \frac{d\mathbf{v}}{dt} = -\mathrm{grad}P + \rho \mathrm{grad}\Phi \tag{2.54}$$

$$\frac{d\rho}{dt} + \rho \mathrm{div}\mathbf{v} = 0 \tag{2.55}$$

$$\frac{1}{P}\frac{dP}{dt} = \frac{\Gamma}{\rho}\frac{d\rho}{dt} \tag{2.56}$$

$$\nabla^2 \Phi = -4\pi G\rho \tag{2.57}$$

The first equation is the equation of motion, the second the equation of continuity, the third the adiabatic equation and the last is the Poisson equation. Φ denotes the gravitational potential and $\mathbf{v}$ is the fluid velocity, Γ is an effective ratio of specific heats ($\rho dP/Pd\rho$) which reduces to γ when γ is constant. The time derivative follows the motion of the fluid. It is related to the derivative at a fixed point by $d/dt = \partial/\partial t + \mathbf{v}\mathrm{grad}$. In an equilibrium situation:

$$\rho = \rho_0(r) \qquad P = P_0(r) \qquad \Phi = \Phi_0(r) \qquad \mathbf{v} = \mathbf{0} \tag{2.58}$$

Now we consider small disturbances about this equilibrium in which the perturbed quantities are functions of all the spatial coordinates and the time. In the equilibrium there is no dependence on spherical polar coordinates. For any variable f we can write:

$$f = f_0 + f_1 \qquad f_1 = Re[\exp(i\omega_{nl}t)\bar{f}_1(r)Y_l^m(\Theta,\Phi)] \tag{2.59}$$

The spherical harmonic is given by:

$$Y_l^m(\Theta,\phi) = P_l^m(\Theta)\exp(im\phi) \tag{2.60}$$

If the star is spherical the oscillation frequency does not depend on m. For the Sun, the departure from sphericity is small and the real oscillation modes have a

behavior close to that shown above but with different m modes having different frequencies. The oscillation frequency ω depends on n and l. The three numbers n, l, m are related to the numbers of times f_1 vanishes in the radial-, Θ- and ϕ-directions and $m \leq l$.

The functions $f_1(r)$ satisfy a system of differential equations and the boundary conditions have to be defined. Since stars do not have sharp surfaces we may assume to a first approximation that all waves are totally reflected at the surface which is defined as the level where density and pressure vanish. A further simplification arises when the change in the gravitational potential produced by the oscillations is unimportant; for most perturbations this is a good approximation because some parts of the star are moving inwards and others moving outwards. We define a perturbation vector ξ by

$$\mathbf{v} = d\xi/dt \tag{2.61}$$

If c_s denotes the velocity of sound in the unperturbed star:

$$c_s = \sqrt{\Gamma P_0/\rho_0} \tag{2.62}$$

one can write:

$$\Psi = c_s^2 \rho_0^{1/2} div\xi \tag{2.63}$$

and the equation for the radial part of ψ is

$$\frac{d^2\psi}{dr^2} = -\frac{1}{c_s^2}\left[\omega^2 - \omega_c^2 - S_l^2\left[1 - \frac{N^2}{\omega^2}\right]\right]\psi \tag{2.64}$$

In addition to the frequency ω three frequencies appear:

- acoustic cut-off frequency ω_c

$$\omega_c^2 = (c_s^2/4H_\rho^2)(1 - 2dH_\rho 2dr) \tag{2.65}$$

  Here, $H_\rho = \rho(d\rho/dr)$ denotes the density scale height,

- Lamb frequency S_l

$$S_l = c_s[l(l+1)]^{1/2}/r \tag{2.66}$$

- Brunt-Väissälä frequency N

$$N^2 = g\left[\frac{1}{\Gamma P}\frac{dP}{dr} - \frac{1}{\rho}\frac{d\rho}{dr}\right] \tag{2.67}$$

  where $g = GM/r^2$.

S_l is always real but ω_c and N can be imaginary. It can be shown that convection occurs when N^2 is negative. We can write our differential equation for ψ as:

$$\frac{d^2\psi}{dr^2} + K_r^2\psi = 0 \tag{2.68}$$

For positive K_r^2 there is a sinusoidal behavior with radius.

For negative K_r^2 we have an exponential dependence giving an exponentially decaying mode which is also called evanescent mode.

In reality K_r depends on r and ω. For different values of ω there are regions in the star where the wave propagates and others where it is evanescent. For both, the high frequency range and the low frequency range $4K_r^2$ is positive: for the high frequency range $\omega > S_l, \omega_c$ and pressure fluctuations are most important; these are the p-modes. For low frequencies $\omega < N$ the g-modes, where the gravity is the restoring force, result.

As it has been already stated, convection occurs where N becomes imaginary. The p-modes can propagate inside the Sun in a region whose lower boundary is determined by the Lamb frequency and whose upper boundary is given by the acoustic cut-off frequency. The g-modes are trapped beneath the convection zone.

2.5.4 Helioseismology and Internal Rotation

The rotation of the Sun can be determined quite straightforward: on the one hand tracers such as sunspots or other phenomena visible on the disk can be used, on the other hand, spectroscopic measurements of the plasma can be used. It was found that the Sun does not rotate like a solid body. It rotates faster at the equator (25 days) and slower near the poles (33 days). Moreover, the rotation rate of sunspots at mid-latitudes is somewhat faster than that deduced from Doppler shifts of the surface plasma.

Our Sun is a middle aged star. The surface rotation rates of young solar-type stars are up to 50 times that of the Sun. Our Sun has lost angular momentum through the magnetized solar wind. Therefore, the outer convection zone must have been gradually spinning down. This also had led to the suggestion that the Sun might still posses a rapidly rotating core, perhaps highly magnetized which also could explain the neutrino problem.

It is extremely important to know the internal rotation of the Sun because the interplay between turbulent motions and rotation with magnetic fields is essential for the solar dynamo which leads to the observed 22 year cycles of magnetic activity. In a spherically symmetric star the frequencies depend upon n and l but not on m. For each (n, l) pair, there is a $(2l + 1)$ fold degeneracy. Rotation breaks the spherical symmetry and lifts the degeneracy. Advection causes a wave propagation with the Sun's rotation to have a higher measured frequency than a similar wave propagating against rotation. Thus the difference in frequency of a pair of oppositely propagating modes is proportional to m times a weighted average of the rotation rate $\Omega(r, \theta)$ where the modes have appreciable amplitude. Here, $\Omega(r, \theta)$ denotes rotation at radius r and latitude θ. The resulting frequency splitting $\Delta\nu_{nlm}$ is half the value of this difference.

Results on the study of the internal solar rotation rate from the SOHO/MIDI instrument are given in Fig. 2.14.

The main results are:

- Differential rotation: occurs only in the convection zone.

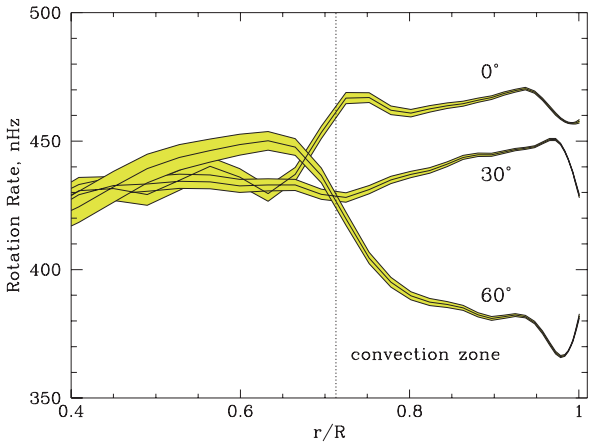


Figure 2.14: This diagram shows the solar rotation rate inferred from two months of MDI Medium-l data as a function of radius at three latitudes, 0 degrees, 30 degrees, and 60 degrees.

- Radiative interior: rotates almost rigidly.
- Thin shear layer near the surface.
- The transition layer between the radiative and convection zone which is called the *tachocline* is mostly located in the radiative zone and thin at the equator but maybe wider at high latitudes.
- There is a sharp radial gradient of the angular velocity beneath the convection zone and the narrow peak of the sound speed at $0.67\,R_\odot$ is due to rotationally turbulent mixing in the tachocline.

More details about these results can be found in Kosovichev *et al.* (1998) [171].

Helioseismology can be used also to give arguments in the question of solar neutrinos. Turck - Chièze *et al.* (2001) [317] used sound-speed and density profiles inferred from SOHO/GOLF and SOHO/ MDI data including these modes, together with recent improvements to stellar model computations, to build a spherically symmetric seismically adjusted model in agreement with the observations. Their model is in hydrostatic and thermal balance and produces the present observed luminosity. Some fundamental ingredients were adjusted, well within the commonly estimated errors, such as the p-p reaction rate ($\pm 1\%$) and the heavy-element abundance ($\pm 3.5\%$); the sensitivity of the density profile to the nuclear reaction rates was examined. The corresponding emitted neutrino fluxes demonstrate that it is unlikely that the deficit of the neutrino fluxes measured on Earth can be explained by a spherically symmetric classical model without neutrino flavor transitions.

New insight into the internal structure of the Sun can be obtained by using time-distance helioseismology. Let us explain this technique by considering seismology on earth. Here, the arrival time of the initial onset of a disturbance is measured. If we know the variation of seismic velocity with depth within the earth, then we can calculate the travel time of rays between an earthquake and a receiver using geometrical approximations. So in principle, we can locate any earthquake in both time and space by recording the arrival times of waves at stations worldwide.

In time-distance helioseismology, the travel time of acoustic waves is measured between various points on the solar surface. To some approximation the waves can be considered to follow ray paths; these depend on a mean solar model. The curvature of the ray paths is caused by increasing sound speed with depth below the surface (see Fig. 2.11). The travel time is affected by various inhomogeneities along the ray path, including flows, temperature inhomogeneities and magnetic fields. The technique consists of a measurement of a large number of times between different locations. Then an inversion method is used to construct 3-D maps of the subsurface inhomogeneities. A review article on that technique was given by Duvall *et al.* (1997) [84].

Inversion Techniques

As we have explained above, the observed oscillation frequencies depend on the physical structure of the solar interior, e.g the variation of quantities such as ρ, T with r. If we assume a spherical symmetric sun and ignore rotational splitting, then we can deduce from our model of the solar interior the corresponding oscillations. Alternatively one can regard $T, \rho...$ as unknowns and use the observed frequencies in order to obtain them. This is called the inversion method. The total number of quantities that can be determined in such a way is equal to the number of observed oscillations. If more frequencies can be identified, a better model of the internal structure can be obtained.

The Seismic Structure of the Sun from GONG data is described in Gough *et al.* (1996) [119].

Solar like oscillations found on other stars are discussed recently e.g. in Bedding and Kjeldsen, 2006 [31] and Kjeldsen *et al.*, 2005 [165] where 37 oscillation modes on α Cen B were found with l=0-3.

Chapter 3

The Solar Atmosphere and Active Regions

3.1 Introduction

The different layers of the Sun and its atmosphere can be defined as follows:

- Solar interior: can be further subdivided into
  1. Core: about 1/3 of the solar radius; here energy production occurs.
  2. Radiation zone: about 1/3 of the solar radius; the energy is transported outward by innumerable emission and absorption processes transferring the high energy γ photons that are produced by nuclear fusion into longer wave photons.
  3. Convection zone: starts below the surface extending about 2×10^5 km into the interior.

- Solar atmosphere: can be subdivided into
  1. Photosphere: starts at the surface [1] and extends up to 500 km.
  2. Chromosphere: above the photosphere; extends to about 2 Mm.
  3. Transition Region: strong increase of temperature up to 10^6 K over a very small spatial range (some 10^4 km).
  4. Corona: starts above 2 Mm, high temperature $> 10^6$ K.

In Fig. 3.1 the variation of temperature and electron density is shown.

[1] which can be defined as the region where light is absorbed considerably over a short distance

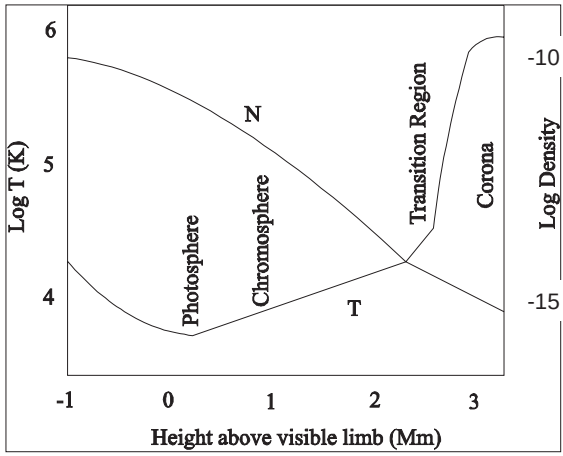


Figure 3.1: Variation of electron temperature and electron density in the solar atmosphere

3.2 Phenomena in the Solar Photosphere

3.2.1 Radiation Transport

The photosphere of the Sun (or of a star) is the layer which can be seen in the visible portion of the continuous radiation spectrum. Here, the photons in the continuum of the visible spectrum have their last scattering encounter before leaving the atmosphere. Let the opacity κ_ν be that fraction of a beam of radiation of frequency ν and intensity I_ν which is absorbed or scattered out of the beam per unit distance. The scattering occurs by atoms, molecules or electrons of the plasma through which it passes. Let us define for an element of plasma of thickness dz and opacity $\kappa_\nu(z)$ the *optical thickness* $d\tau_\nu$ (the subscript ν denotes that this quantity depends on the frequency)[2] by:

$$d\tau_\nu = -\kappa_\nu(z)dz \tag{3.1}$$

and hence

$$\tau_\nu(z) = -\int_0^z \kappa_\nu(z)dz \tag{3.2}$$

The transfer of radiation through the atmosphere of a star is governed by the equation of radiative transfer. If θ denotes the angle between the direction of the beam of radiation and the outward normal, and $\mu = \cos\theta$, then under the assumptions that a) the atmosphere is plane - parallel and b) is locally in thermodynamic equilibrium (LTE), the transport equation becomes:

$$\mu\frac{\partial I_\nu(\tau_\nu,\mu)}{\partial\tau_\nu} = B_\nu(T) - I_\nu(\tau_\nu,\mu) \tag{3.3}$$

[2]Very often the solar surface is defined as the layer where $\tau_{500\,\mathrm{nm}} = 1$

where $B_\nu(T)$ is the Planck function at temperature T:

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\left(e^{h\nu/kT} - 1\right)^{-1} \tag{3.4}$$

An elementary solution yields for the intensity of radiation emerging in direction μ:

$$I_\nu(\mu) = \int_0^\infty B_\nu(T) e^{-\frac{\tau_\nu}{\mu}} \frac{d\tau_\nu}{\mu} \tag{3.5}$$

The Planck function must increase with depth, since the temperature increases with depth (see Fig. 3.1). Eddington made the following Ansatz assuming a linear increase of the function B_ν with depth:

$$B_\nu = C + D\tau_\nu \tag{3.6}$$

If we put this into 3.5, we arrive at

$$I_\nu = C + D\mu \tag{3.7}$$

The physical depth z corresponding to $\tau_\nu = 1$ is said to be the origin of the emergent radiation of frequency ν. Thus, by observing the photosphere at different frequencies, we sample it at different heights. Since τ_ν is related to the absorption coefficient, the variation of κ_ν defines how deep we look into the atmosphere at a given frequency ν.

For the Sun and solar like stars, the continuum absorption coefficient is formed by the negative H ion H^-.

The deepest penetration is obtained at IR wavelengths (about 1.6 μm); higher layers may be sampled by observing at the centers of absorption lines. The greater the optical depth at a given wavelength the less radiation reaches the observer from that layer.

If we look at the solar disk we immediately see that the central regions are brighter than the limb. The function $I_\nu(\mu)/I_\nu(1)$ is called the *limb darkening* (center to limb variation). This may be written as:

$$\frac{I_\nu(\mu)}{I_\nu(1)} = \int_0^\infty \frac{B_\nu(T)}{I_\nu(1)} e^{-\tau_\nu/\mu} \frac{d\tau_\nu}{\mu} \tag{3.8}$$

If one does an inversion of this equation information about the physical structure (temperature distribution) of the solar atmosphere is obtained. Stellar limb functions can not be measured accurately so this method is only applicable to the Sun.

3.2.2 Granulation

Under very good seeing conditions the Sun shows a cellular like pattern which is called granulation. The mean diameter of the cells is about 1000 km which corresponds roughly to 1 arcsec (as seen from the Earth). In the bright granules matter is streaming upwards, in the darker intergranular lanes streaming downwards. Up

to now the best granulation images have been taken from the ground since no large solar telescopes have been launched. In 2006 SOLAR B will be launched. This will be the first large optical telescope flown in space. Its aperture is 50 cm and angular resolution achieved will be 0.25 arcsec.

In order to minimize the effect of the turbulence of the Earth's atmosphere (seeing), the exposure times must be shorter than 1/10 s. Usually, one makes a burst of several images and then selects the best image for further analysis. Spectrograms show a high degree of correlation between intensities and velocities proving the convective character of the phenomenon. Under a spatial resolution better than 0.5 arcsec, the situation becomes more complex. Regular granules seem to have a maximum for the upflow near their center, so called exploding granules have a maximum upflow between the center and the edge. Measuring the width of spectral lines one gets a hint for turbulence. Enhanced line widths indicate enhanced turbulence. It was found that turbulence is located in the downdrafts which is also predicted by 3 D models. The turbulence may be generated by the shear between upflows and downflows at granular borders and on transonic flows. A review about solar granulation was given by Muller (1999) [227] where further references can be found.

A problem to investigate the granulation is how can we identify a granulum? One possibility is to identify them by an isophote contour at a level close to the average intensity of the photosphere. The images must be filtered in order to remove the intensity fluctuations at low frequency, originating in instrumental brightness inhomogeneities and in solar large scale fluctuations (which arise from the supergranulation, mesogranulation and oscillations). Finally, high frequency noise must be eliminated. In the Fourier domain such a filter has the form:

$$F(k) = (1 - e^{-Ca_1^2k^2})e^{Ca_2^2k^2} \tag{3.9}$$

The parameters are chosen, so that the maximum filter transmission stays in between spatial scales 0.5 and 1 arcsec. Such a filter is partially restoring as it enhances the contrast of the smallest granules which can then be identified more clearly. Another method is to find the inflection points of the intensity distribution in the image using a Laplacian operator.

How do granules evolve? The most common process is that of fragmentation: a granule grows and then splits into several fragments (3-4). About 60% of granules appear or die by this process. Some granules appear spontaneously in intergranular spaces and grow, others result from merging of two adjacent granules. The most spectacular evolution is observed for exploding granules. The granule lifetime can be determined by their visual identification on successive images or by cross correlating these images. There is a large discrepancy of the results: granular lifetimes range from 6 to 16 min.

From the physical point of view, there exists a limitation for the horizontal expansion because of mass conservation and radiative loss. Matter is streaming upward in a granulum, expands and horizontal flows are driven by pressure gradients; thus the central upflow is decelerated which then cannot supply the horizontal expansion and the radiative loss. The central part cools and the granule splits into several fragments, after a downdraft developed. On the other hand, intergranu-

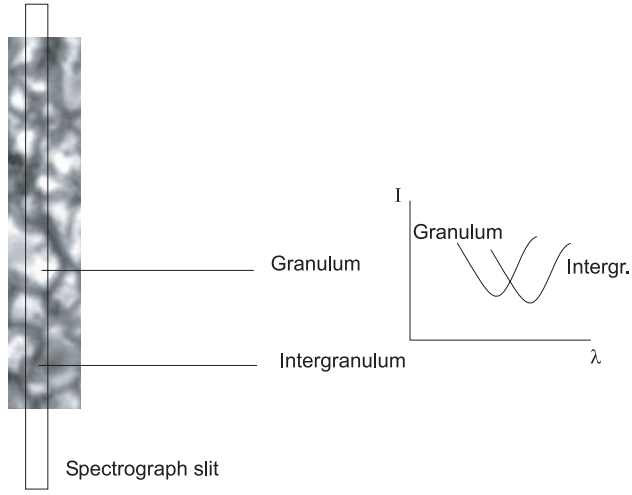


Figure 3.2: Spectroscopic observation of solar granulation. The entrance of a spectrograph slit covers different granular/intergranular areas. Line profiles emanating from granules are blueshifted because matter moves upwards and profiles from intergranular areas are redshifted because matter moves away from the observer. This is valid for solar granulation observed near the disk center.

lar lanes are interconnected without interruption. They contain some dark holes which exist over 45 min and may correspond to the fingers of downflowing material predicted by 3 D models.

Using time series with the 50 cm refractor at the turret dome of the Pic du Midi observatory Roudier *et al.* (1997) [262] showed the existence of singularities in the intergranular lanes what they called intergranular holes which have diameters between 0.24 arcsec and 0.45 arcsec and are visible for more than 45 min. These holes appear to be systematically distributed at the periphery of mesogranular and supergranular cells.

Spectroscopic observations of the solar granulation with high resolution yield information about velocities e.g. when observed near solar disk center, granular profiles are blueshifted because matter rises and moves in direction to the observer (see Fig. 3.2).

Concerning the structural properties of granules, we have to mention that their number N increases monotonically with decreasing size. Granules of size 1.4 arcsec are the main contributors to the total granule area. When the area A is plotted versus their perimeter in a log-log scale, the dispersion of points (each of them marks a granule) is small and their shape can be characterized by the relation:

$$P \sim A^{D/2} \tag{3.10}$$

where D is the fractal dimension. It seems that there are two ranges with different fractal dimensions:

- $D \sim 1.25$ for granules smaller than about 1.35 arcsec.
- $D \sim 2.00$ for granules that are larger than 1.35 arcsec.

The physical interpretation is as follows: In hydrodynamics, the fractal dimension is often used to get some information about the dynamical state. In the theory of Kolmogorov (he treated isotropic, homogeneous turbulence in three dimensions and obtained a 5/3 power law for the energy spectrum) a value of $D = 5/3$ is predicted for isotherms and 4/3 for isobars. A fractal dimension of 2 or even larger means that the shape is complex which is confirmed by observations since many of them are in the process of fragmentation.

Granules above 1.4 arcsec have nearly the same brightness, the intergranular brightness is nearly constant, with an average value of 0.92 (when the averaged continuum is at 1.0). The rms intensity fluctuations of the best image is 10-11% at $\lambda\,465,\mathrm{nm}$ (50 cm refractor at La Palma) and 8-9% at $\lambda\,570\,\mathrm{nm}$ (50 cm refractor at the Pic du Midi). Restored values lie between 10 and 22%. From the granular contrast we can infer the temperature variations (assuming Planck's law) which correspond to $\sim 200\,\mathrm{K}$.

Theoretical Approaches

The simplest model of convection is the classical Rayleigh problem: suppose a fluid (either gaseous or liquid), confined between two horizontal plates separated by a distance h and maintained at temperature T_1 (upper) and T_2 (lower) with $T_2 > T_1$. If the fluid has a positive coefficient of thermal expansion α as it will be the case for a gas and for a normal fluid, the fluid near the lower plate will tend to rise. However, this will be opposed by two effects: a) viscous dissipation, b) thermal diffusion in the fluid. Convection will occur when the imposed temperature gradient $(T_2 - T_1)/h$ is sufficiently large or, for a given gradient, when the coefficients of the kinematic viscosity ν and of thermal diffusion κ are sufficiently small. Rayleigh's theoretical analysis of the problem in 1916 inspired Bénard to investigate this 40 years later. It was found that convective instability occurs when the Rayleigh number R exceeds a critical value:

$$R > R_{\mathrm{crit}} \qquad R = \frac{g\alpha\beta h^4}{\kappa\nu} \tag{3.11}$$

where β is the temperature gradient. For R_{crit} Rayleigh found the value 657.5. This value depends on the boundary conditions. Later Chandrasekhar has shown that e.g. a Coriolis force (as an effect of rotation) inhibits the onset of instability to an extend which depends on the value of a non dimensional parameter (called Taylor number):

$$C = \frac{4h^4\Omega^2}{\nu^2} \tag{3.12}$$

here, Ω is the vertical component of the angular velocity vector. For details see e.g. Chandrasekhar (1961) [62].

For the solar convection zone R is extremely high, $R \sim 10^{10\ldots11}$.

Important information about the origin of the solar granulation can be inferred from power spectra. From spectrograms we can obtain 1-D power spectra of intensity and velocity fluctuations, from white light images, one gets 2-D power spectra for the intensity fluctuations. The theoretical power spectrum of the velocity fluctuations decreases as $k^{-5/3}$ down to the scale of molecular diffusion. The temperature power spectrum however decreases as $k^{-5/3}$ only to a scale k_c. At smaller scales the spectrum decreases as $k^{-17/3}$(Espagnet *et al.*, 1995 [90]). Thus k_c separates the inertial convective range, where heat advection dominates from the inertial conductive range, where diffusion dominates. The former is the range of large granules, the latter the range of small granules.

The basic set of hydrodynamic equations to describe solar convection is described in detail in Nordlund, 1982 [234].

Interaction between Granulation and Magnetic Elements

In this section we consider magnetic regions which occur as Plages or faculae (in active regions) and in the photospheric network (in the quiet Sun) in the form of small bright points. Sunspots will be discussed in the next paragraph. Magnetic elements (observed in high resolution magnetograms) and bright points (observed in high resolution filtergrams) coincide. Bright points are visible in white light near the limb (e.g. as faculae) but also at the disk center because they have a brightness comparable to granules. It is very easy to observe them with a G Band filter (see e.g. Kiselman *et al.* 2001 [164]). Fraunhofer (1817) denoted a roughly 1 nm wide band with CH lines around $\lambda = 430.5nm$ by G in his initial inventory of the visible solar spectrum. This region is a principal diagnostic to study photospheric magnetism at the highest achievable angular resolution (Muller *et al.*, 1989 [228]).

The dynamics of the granules forces these small bright points to appear and stay in the intergranulum when the surrounding granules converge. Thus there seems to be a continuous interaction between granules and magnetic elements. Small magnetic flux tubes are the channels along which the energy is carried in upper layers by different kinds of waves. In that context Choudhuri *et al.* (1993) [65] discussed the generation of magnetic kink waves by rapid footpoint motions of the magnetic flux tubes. They found that these pulses are most efficient. Kalkofen (1997) [152] discussed the impulsive generation of transverse magneto-acoustic waves in the photosphere, propagating upward with exponential growth of amplitude. Such waves are observed as intensity oscillations in the H and K lines of Ca II in network bright points.

Granulation-Mesogranulation

Idealized numerical experiments on turbulent convection were made by Cattaneo *et al.* (2001) [58]. The authors found two distinct cellular patterns at the surface. Energy-transporting convection cells (corresponding to granules in the solar photosphere) have diameters comparable to the layer depth, while macrocells (corresponding to mesogranules) are several times larger. The motion acts as a small-scale turbulent dynamo, generating a disordered magnetic field that is concentrated at macrocellular corners and, to a lesser extent, in the lanes that

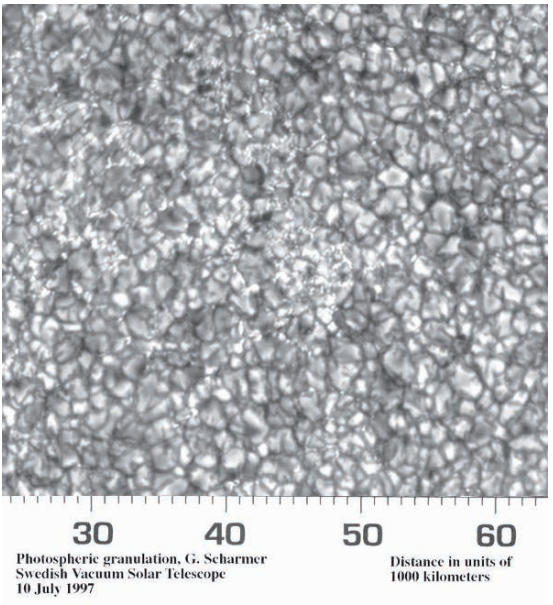


Figure 3.3: Solar granulation and small network bright points

join them. These results imply that mesogranules owe their origin to collective interactions between the granules.

3.2.3 Five Minutes Oscillations

In 1962 Leighton, Noyes and Simon [192] identified a strong oscillatory component which they called five minutes oscillations because of its characteristic period. Later, these were interpreted as standing acoustic waves trapped in resonant cavities below the photosphere.

The spatial relation between the 5-min oscillations and the granulation pattern has been largely debated in the literature. Of course such a discussion is important to understand the excitation mechanism of these oscillations and, hence, the internal properties of the Sun. Theoretical studies suggest that acoustic waves which comprise the 5-min oscillations are stochastically generated by turbulent convection just beneath the photosphere (Goldreich *et al.*, 1994 [115]). Espagnet *et al.* (1996) [91] studied the relation between oscillation and granulation and found that the most energetic oscillations are concentrated in downflow regions in expanding intergranular spaces. This was later confirmed by Goode *et al.* (1998) [116].

Strous *et al.* ((2000) [302]) found a roughly linear relation between the peak seismic flux and the peak downward convective velocity associated with each seismic event.

Other authors like e.g. Hoekzema *et al.*, 1998 [135], who analyzed G band images found that photospheric 5 min oscillations are global and rather insensitive to local fine structure.

Using a 30-min time series of CCD spectrographs, Khomenko, Kostik and Shchukina, 2001 [160], found different amplitudes, phases and periods of the 5-min oscillations above granules and intergranular lanes. The most energetic intensity oscillations occurred above intergranular lanes, the most energetic velocity oscillations above granules and lanes with maximum contrast that are cospatial with regions with maximum convective velocities.

3.2.4 Sunspots

Discovery of Sunspots

When the Sun is very low just above the horizon one can make a short glimpse on it with the unprotected naked eye. Chinese astronomers were the first who reported on dark spots visible on the Sun. In the year 1611 sunspots were observed for the first time through a telescope by four men: J. Goldsmid (Holland), G. Galilei (Italy), Ch. Scheiner (Germany) and Th. Harriot (England). The first publication on that topic appeared from Goldsmid (he is better known by his Latin name Fabricius). He even argued that the Sun must rotate since the sunspots move across the disk. Since he was a Jesuit he first suspected some defect in his telescope when he observed the spots. Then he failed to persuade his ecclesiastical superiors who refused to allow him to publish his discovery. However, Scheiner announced his discovery in three anonymous letters to a friend of Galileo and Galileo responded in three letters in 1612 (the sunspot letters) that he had discovered the sunspots. Of course Scheiner and Galileo became enemies. Scheiner later reported his discoveries in his work Rosa Ursinae sive Sol in 1630. Both scientists noted that the spots appear only within zones of low latitudes at either side of the equator. There are never spots near the poles.

After the initial interest and the publication of Scheiner's major work the interest in sunspots vanished. In 1977 Eddy showed that this must be seen in connection with the fact that during 1640-1705 there was a great reduction in the number of sunspots seen on the Sun which is now known as the *Maunder Minimum*.

The next significant discovery was made by Schwabe who was a German apothecary and bought a telescope in 1826 in order to search for a planet inside the orbit of Mercury. He recorded the occurrence of sunspots over 43 years and reported on a periodicity of their occurrence of about 10 years. In 1851 appeared his publication on the 11 year periodicity of the annually averaged sunspot numbers. Several years later Carrington showed from his observations that the Sun rotates differentially; a point at the equator rotates more rapidly than one at higher latitudes. He defined an arbitrary reference point on latitude 10^0 as longitude zero and a rotation completed by this point is known as Carrington rotation (CR)[3]. The sideral Carrington rotation is 25.38 days, the synodic value varies a

[3]For example on March 14 2006 Carrington Rotation 2041 started at 14.43 UT and ended on April 10, at 21.47 UT

little during the year because of the eccentricity of the Earth's orbit (its mean value is 27.2753 days).

Carrington was also the first to see a white light flare on the Sun in the morning of Sep. 1, 1859, during sketching sunspot projections with a friend. Suddenly two crescent-shaped patches broke out, brightened, moved a distance twice their length, then faded away as two dots within five minutes. Carrington reported to the Royal Astronomical Society that at 4 hours after midnight the magnetic instruments indicated a great magnetic storm. So he was in fact the first who noticed that there exists a connection between solar phenomena and disturbances on Earth.

R. Wolf (1816-1893) studied all available records and derived a more accurate estimate for the sunspot cycle. In 1848 he introduced the relative (Zurich) sunspot number R_Z as a measure for solar activity. Sunspot often appear as groups. If g denotes the number of sunspot groups and f the number of individual spots, then

$$R_Z = k(10g + f) \tag{3.13}$$

k... personal reduction factor. Today more than 30 observatories contribute to determine this value.

The Physics of Sunspots

Sunspots consist of dark central regions, called umbra and a surrounding less dark filamentary region called penumbra. The umbral diameter is about 10 000 km but for the largest spots may exceed 20 000 km. Penumbral diameters are in the range of 10 000 -15 000 km. Sunspots evolve and some of them are visible over more than 1 rotation period. The observations of sunspots showed that the rotation of the Sun is not like that of a solid body.

Another interesting phenomenon is the *Wilson depression*. In 1769 Wilson observed a very large spot nearing the west limb and noted that the penumbra on the further side from the limb gradually contracted and finally disappeared. When the spot reappeared at the east limb some two weeks later, the same behavior was displayed by the penumbra on the opposite site of the spot. The surface of a sunspot is depressed below the surface of the surrounding plasma.

The temperature of the umbra is about 4 000 K whereas the temperature of the solar surface is about 6 000 K. According to Stefan's law the total energy emitted per unit area by a black body at temperature T is proportional to T^4; the above mentioned temperature difference between umbra and photosphere means that the energy flux through a given area of the umbra is $\sim 20\%$ of that through an equivalent area of the photosphere. The penumbra has a temperature between umbra and solar surface. In the penumbra we observe also a radial outflow of matter with the velocity increasing outwards with a characteristic speed of 1 to 2 kms/s (*Evershed effect*).

In 1908 Hale discovered that the spectral lines are split in the sunspots. This is caused by the *Zeeman effect* in the presence of strong magnetic fields. In the absence of magnetic fields several quantum mechanical state may possess the same energy but the magnetic fields destroy this symmetry resulting in a splitting of

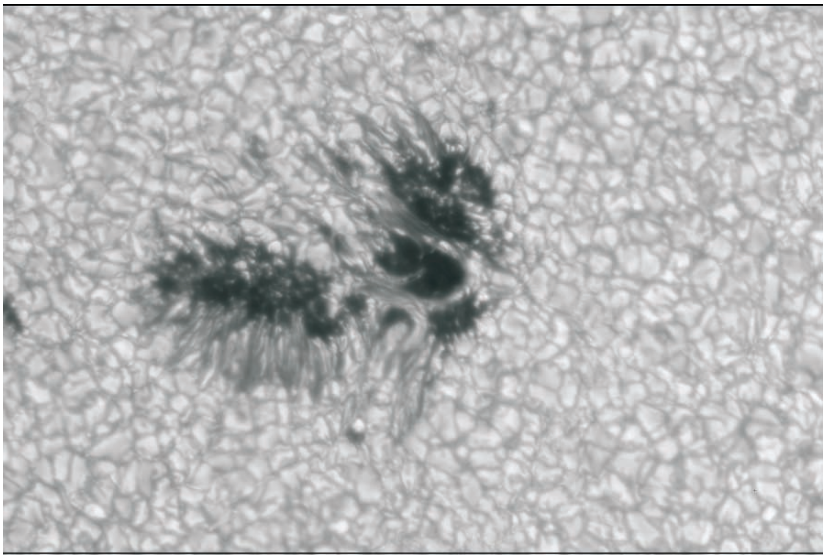


Figure 3.4: Large sunspot showing the dark central umbra and the filamentary penumbra. Outside the penumbra the granulation pattern is clearly seen. Courtesy: M. Sobotka, A.H., SST, La Palma, 2003

the energy levels. The displacement of the lines due to the Zeeman effect is given by:

$$\Delta\lambda = 4.7 \times 10^{-8} g^* \lambda^2 B \tag{3.14}$$

The wavelength λ is given in nm, the Landéfactor g^* depends on the spin and orbital momentum of the levels and B denotes the magnetic induction given in Tesla.

$$1\,\text{Tesla} = 10^4 \text{Gauss} = \text{Vs/m}^2 \tag{3.15}$$

The strength of the magnetic field is in the order of 3 000 Gauss.

Small dark spots with diameters $< 2\,500$ km lacking penumbrae are called pores. They exist within groups or appear also as isolated structures. Their lifetimes are in the range of a few hours to several days.

Sunspot groups tend to emerge either sequentially at the same or similar Carrington longitudes, which are designated as active longitudes, or to overlap in clusters. The distribution of sunspots is non-axisymmetric and spot group formation implies the existence of two persistent active longitudes separated by 180^o Usoskin, Berdyugina, Poutanen, 2005 [322].

High Spatial Observations of Spots

High spatial resolution observations of sunspots show that there appear a lot of different morphological phenomena: multiple umbrae, bright umbral dots, light

bridges, dark nuclei in the umbra etc. One problem in the study of sunspots and their fine structure is observational stray light.

An important photometric parameter of umbral cores is the minimum intensity (intensity of the darkest point) $I_{\min}$ which is usually in the range of 0.05-0.3 of the mean photospheric intensity at $\lambda \sim 540\,\mathrm{nm}$. There seems to be a relation between the size of the umbrae and the temperature. Umbrae with a diameter $D_U < 7''$ have higher temperatures than the large ones. Moreover, regions with higher magnetic field strength are darker and cooler than those with lower strength. The darkest regions in umbral cores are dark nuclei. These are the areas with the strongest magnetic fields and the orientation of the field is perpendicular to the surface of the Sun. They are not necessarily centered in the umbral cores, some of them are observed close to the edge of the penumbra. They cover 10-20% of the total umbral core area and their size is about 1.5 arcsec[4]. The penumbra shows elongated structures which is a consequence of the strongly inclined magnetic field. Bright penumbral filaments consist of penumbral grains. They seem to have cometary like shapes with "heads" pointing towards the umbra and have a mean width of only $0.36''$ and a length of $0.5...2''$. The observed brightness approaches the photospheric one and the lifetimes are between 40 minutes and 4 hours. They are separated by narrow dark fibrils. The magnetic field seems to be stronger and more horizontal in dark fibrils and weaker and more vertical in penumbral grains.

It is also interesting to note that nearly all penumbral fine structures are in motion. The penumbral grains move towards the umbra with an average speed of 0.3-0.5 km/s. On the other hand, dark cloud like features which arise from the dark fibrils move rapidly outwards (up to 3.5 km/s) towards the outer penumbral border.

The last fine structure which is important to study are the light bridges . They cross the umbra or penetrate deeply into it and can be observed for several days although they change their shape substantially on the scale of hours. They can be classified into faint (located inside umbral cores) and strong (separating umbral cores). Strong light bridges separate umbral cores of equal magnetic polarities and a subclass of them opposite polarities. The analysis of 2-D power spectra of intensity fluctuations inside strong light bridges showed that the "granules" that can be seen there are smaller (1.2 arcsec, normal granulation: 1.5 arcsec) and the slopes of power spectra indicated the presence of a Kolmogorov turbulent cascade. The magnetic field strength in strong light bridges is substantially lower than in adjacent umbra.

A recent review about the fine structure of sunspots was given by Sobotka (1999) [290] where other references can be found. A review on empirical modelling and thermal structure of sunspots was given by Solanki (1997) [292].

Sunspots and Magnetic Fields

Observations demonstrated, that spots often occur in bipolar magnetic groups. The magnetic polarity of the leading spot in the pairs (in terms of solar rotation) changes from one 11 year cycle to the next- this is known as *Hale's law*. There is a

[4] $1''$=1 arcsec corresponds to about 750 km on the solar surface

22 year magnetic cycle. Spots appear as a magnetic flux tube rises (see magnetic buoyancy) and intersects with the photosphere. The magnitude of the magnetic induction is 0.3 T in the umbra and 0.15 T in the penumbra. In the umbra the field is approximately vertical, and the inclination increases through the penumbra.

Hale's observations also suggested that the Sun has an overall dipolar magnetic field (10^{-4} T). This very weak dipolar field is reversed over the magnetic cycle. Almost all of the photospheric field outside sunspots is concentrated in small magnetic elements with a magnetic induction between 0.1 and 0.15 T.

Only the surface properties of the flux tube that defines a spot can be observed. The question is, how the field structure changes with depth. The simplest model is a monolithic column of flux. Let us assume that the pressure inside the flux tube is negligible compared to the magnetic pressure. We also assume that the gravitational force is unimportant in obtaining an approximate idea of the magnetic field structure, the magnetic field in cylindrical polar coordinates can be taken to be current free:

$$\mathbf{B} = \frac{1}{\omega}\left[-\frac{\mathrm{d}\psi}{\mathrm{d}z}, 0, \frac{\partial\psi}{\partial\omega}\right] \tag{3.16}$$

Thus curl$\mathbf{B} = \mathbf{0}$. Since div$\mathbf{B} = \mathbf{0}$,

$$\frac{\partial^2\psi}{\partial\omega} - \frac{1}{\omega}\frac{\partial\psi}{\partial\omega} + \frac{\partial^2\psi}{\partial z^2} = 0 \tag{3.17}$$

The neighboring photosphere, in which the flux tube is embedded has a known pressure variation with height $P_\mathrm{e}(z)$. The boundary of the flux tube is at $\omega = \omega_0(z)$, where

$$B^2/2\mu_0 = P_\mathrm{e}(z) \tag{3.18}$$

We see that as $z \to \infty$ the field becomes nearly horizontal and $B_\omega \sim F/2\pi\omega_0^2$ and as $z \to -\infty$, the field becomes vertical and $B_\omega \sim F/\pi\omega_0^2$.

There is one problem with this monolithic model: the difference in the energy radiated by the spot and by an equivalent area of the normal photosphere is only about a factor of 4. This is less than would be expected if convection in the spot were completely suppressed. Therefore, it is believed that some form of convective energy transport must occur and the field must be more complex e.g. coherent flux tubes or a tight cluster. Reviews about these topics were given by Bogdan (2000) [40] and Hurlburt (1999) [144].

Using the 1 m Swedish Solar Telescope a high resolution study of the inclination of magnetic fields within sunspots was performed by Langhans *et al.*, 2005 [187]. Within sunspots, dark penumbral cores, and their extensions into the outer penumbra, are prominent features associated with the more horizontal component of the magnetic field from about 40^0 in the inner penumbra to nearly horizontal in the middle penumbra. Bright flux component is associated with a more vertical field component.

Sunspot Group Classification

The 3 component McIntosh classification (McIntosh, 1990) [218] is based on the general form 'Zpc', where 'Z' is the modified Zurich Class, 'p' describes the penum-

bra of the principal spot, and 'c' describes the distribution of spots in the interior of the group. This classification scheme substituted the older scheme that was introduced by Waldmeier (1938).

1. Z-values: (Modified Zurich Sunspot Classification).

   A - A small single unipolar sunspot. Representing either the formative or final stage of evolution.
   B - Bipolar sunspot group with no penumbra on any of the spots.
   C - A bipolar sunspot group. One sunspot must have penumbra.
   D - A bipolar sunspot group with penumbra on both ends of the group. Longitudinal extent does not exceed 10 deg.
   E - A bipolar sunspot group with penumbra on both ends. Longitudinal extent exceeds 10 deg but not 15 deg.
   F - An elongated bipolar sunspot group with penumbra on both ends. Longitudinal extent of penumbra exceeds 15 deg.
   H - A unipolar sunspot group with penumbra.

2. p-values:

   x - no penumbra (group class is A or B)
   r - rudimentary penumbra partially surrounds the largest spot. This penumbra is incomplete, granular rather than filamentary, brighter than mature penumbra, and extends < 3 arcsec from the spot umbra. Rudimentary penumbra may be either in a stage of formation or dissolution.
   s - small, symmetric (like Zurich class J). Largest spot has mature, dark, filamentary penumbra of circular or elliptical shape with little irregularity to the border. The north-south diameter across the penumbra is ≤ 2.5 degrees.
   a - small, asymmetric. Penumbra of the largest spot is irregular in outline and the multiple umbra within it are separated. The north-south diameter across the penumbra is $\leq$ than 2.5 degrees.
   h - large, symmetric (like Zurich class H). Same structure as type 's', but north-south diameter of penumbra is more than 2.5 degrees. Area, therefore, must be larger or equal than 250 millionths solar hemisphere.
   k - large, asymmetric. Same structure as type 'a', but north-south diameter of penumbra is more than 2.5 degrees. Area, therefore, must be larger or equal than 250 millionths solar hemisphere.

3. c-values:

   x - undefined for unipolar groups (class A and H)
   o - open. Few, if any, spots between leader and follower. Interior spots of very small size. Class E and F groups of 'open' category are equivalent to Zurich class G.
   i - intermediate. Numerous spots lie between the leading and following portions of the group, but none of them possesses mature penumbra.
   c - compact. The area between the leading and the following ends of the

spot group is populated with many strong spots; at least one interior spot shows a mature penumbra. The extreme case of compact distribution has the entire spot group enveloped in one continuous penumbral area.

There exists also the Mount Wilson classification scheme:
α: Denotes a unipolar sunspot group.
β: A sunspot group having both positive and negative magnetic polarities, with a simple and distinct division between the polarities.
$\beta - \gamma$: A sunspot group that is bipolar but in which no continuous line can be drawn separating spots of opposite polarities.
δ: A complex magnetic configuration of a solar sunspot group consisting of opposite polarity umbrae within the same penumbra.
γ: A complex active region in which the positive and negative polarities are so irregularly distributed as to prevent classification as a bipolar group.

Sunspots and the Solar Cycle

The number of sunspots changes with a 11 years period which is called the solar activity cycle. Today we know that all solar activity phenomena are related to sunspots and thus to magnetic activity. To measure the solar activity the sunspot numbers were introduced and in order to smear out effects of solar rotation, R is given as a monthly averaged number and called the sunspot relative number. Today there exist better methods to quantify the solar activity however sunspot numbers are available for nearly 400 years and thus this number is still used.

The Royal Greenwich Observatory (RGO) compiled sunspot observations from a small network of observatories to produce a data set of daily observations starting in May of 1874. The observatory concluded this data set in 1976 after the US Air Force (USAF) started compiling data from its own Solar Optical Observing Network (SOON). This work was continued with the help of the US National Oceanic and Atmospheric Administration (NOAA) with much of the same information being compiled through to the present.

Since 1981, the Royal Observatory of Belgium harbors the Sunspot Index Data center (SIDC), the World data center for the Sunspot Index. Recently, the Space Weather forecast center of Paris-Meudon was transferred and added to the activities of the SIDC. Moreover, a complete archive of all images of the SOHO instrument EIT has become available at the SIDC.

Let us briefly summarize the behavior of sunspots during the activity cycle:

- The leader spots (i.e. by convention it is defined that the Sun rotates from east to west; the largest spot of a group tends to be found on the western side and is called the leader, while the second largest in a group is called the follower) in each hemisphere are generally all of one polarity, while the follower spots are of the opposite polarity.

- If the leaders and followers are regarded as magnetic bipoles, the orientation of these bipoles is opposite on opposite hemispheres.

- The magnetic axes of the bipoles are inclined slightly towards the equator, the leader spot being closer. This inclination is about 12^0.

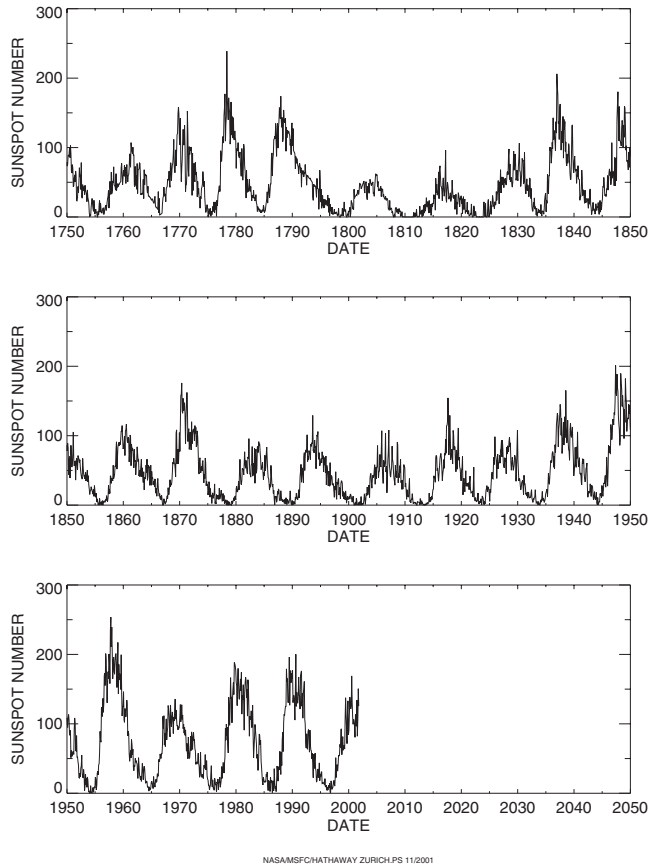


Figure 3.5: Relative Sunspot number. Two cycles can bee seen, the normal cycle with about 11 years and the so called Gleissberg cycle with a period of about 80 years.

- Towards the end of a cycle spot groups appear at high latitudes with reversed polarity, they belong to the new cycle whereas those with normal polarity for the old cycle occur close to the equator. This is illustrated in the so called butterfly diagram (see Fig. 3.6).

In Table 3.1 some parameters for the energetics of large sunspots are given, i.e. spots with a diameter $\geq 3.5 \times 10^4$ km. Penumbral waves are horizontal outwards waves (in H_α) with velocities between 10 and 20 km/s.

3.2.5 Photospheric Faculae

Near the solar limb, regions brighter than the surrounding photosphere can be found and are known as photospheric faculae. These structures are hotter than their surroundings. At the disk center they are not visible. In the neighborhood of

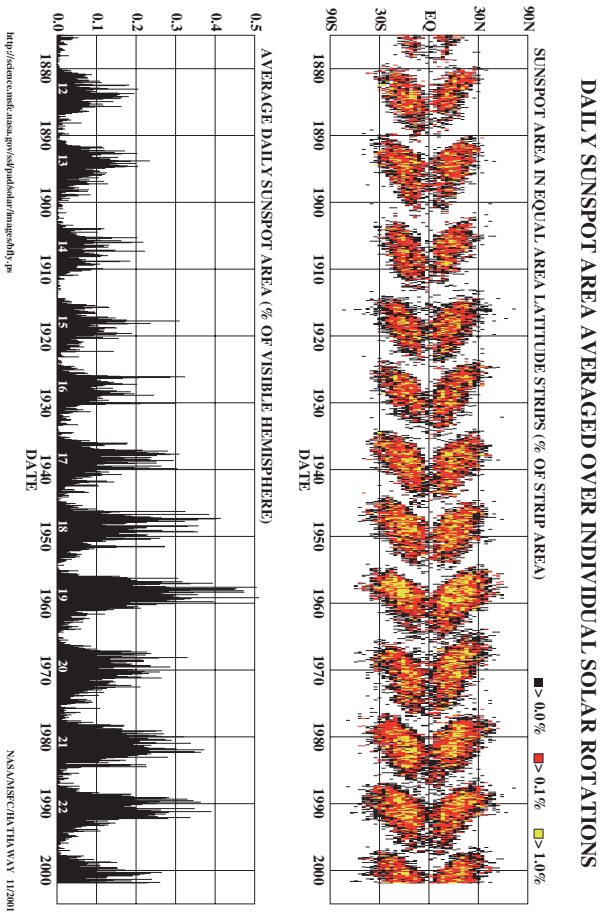


Figure 3.6: Butterflydiagram illustrating the equatorward motion of spots during the activity cycle.

sunspots they tend to overlap and can be identified further from the limb. They appear in increased numbers in a region prior to the emergence of sunspots and remain for a rotation or more after the spots have decayed. As it will be shown later they are important for the energy balance between sunspots and the photosphere. Faculae can be observed on the whole disk using filtergrams . In that case they are often called plage and attributed to the chromosphere. Photospheric faculae are manifestations of concentrated azimuthal magnetic fields. One possibility to study sunspots and faculae at photospheric levels is to use the Ca II K line 0.05 nm off the center with a 0.015 nm passband.

Polar faculae appear as pointlike, bright photospheric spots near the solar limb at latitudes of 55 degrees or more (average of 65 degrees). Polar faculae tend to occur at lower latitudes (as low as 45 degrees) during the years in which there are only few observable. They can be distinguished from main zone faculae by

Table 3.1: Sunspot energy values (from [17])

	$\mathrm{erg\,cm^{-2}\,s^{-1}}$	Total $\mathrm{erg\,s^{-1}}$
missing flux, umbra	4.7×10^{10}	7×10^{28}
missing flux, penumbra	1.2×10^{10}	1×10^{29}
Alfvén waves (umbra)	10^{11}	1×10^{29}
running penumbral waves	3×10^{8}	3×10^{27}

their essentially pointlike and solitary appearance, in contrast to the more area- and grouplike appearance of the main zone faculae (55 degrees or lower). Their lifetime is shorter (minutes to hours) than that for ordinary faculae. The brightest can last for a couple of days, and can be traced farther from the solar limb too. In connection with the activity cycle it is interesting to note that polar faculae are most numerous at times of minimum solar activity, which in turn might be an additional hint for their relation with the upcoming new solar cycle.

3.3 The Chromosphere

3.3.1 Diagnostics

The chromosphere [5] lies between the corona and the photosphere and can be observed during short phases of solar eclipses. The spectrum obtained at these rare occasions is called a flash spectrum. Above the photosphere the temperature passes through a minimum of 4 000 K and then rises to several 10^4 K in the chromosphere and much more rapidly in the transition region until the coronal temperature ($\sim 10^6$ K) is reached. Two very prominent spectral lines formed in the chromosphere are the so called H and K lines of singly ionized Ca (called Ca II). These lines are in absorption in the spectrum of the photosphere but appear as emission lines in the hotter chromosphere. Their strength varies through the sunspot cycle, the lines are stronger at maximum [6]. Important chromospheric lines are listed in Table 3.2, the physics of the formation of these lines is complicated since the assumption of LTE is not valid.

The temperature variation throughout the chromosphere can be described as follows:

- Temperature minimum: near 500 km; here the UV continuum near 160 nm, the far IR continuum and the minima in the wings of Ca II and Mg II lines are formed,
- moderately fast temperature increase from $T_{\min}$ to approx. 6 000 K. In the first plateau there are the emission peaks of Ca II and Mg II, the center of

[5] A classical textbook about the chromosphere is: The Solar Chromosphere and Corona, R.G. Athay, 1976, Reidel

[6] The observations of the variation of the strength of stellar H and K lines provide thus information about stellar activity cycles.

Table 3.2: Prominent chromospheric emission lines

Line	Wavelength
Lyα	121.6 nm
Lyβ	102.6 nm
C I continua	$\leq$ 110.0 nm, $\leq$ 123.9 nm
Mg II h	280.3 nm
Mg II k	279.6 nm
Ca II H	396.8 nm
Ca II K	393.4 nm
He I	447.1 nm, 587.6 nm
Ca II IR	849.8, 854.2, 866.2 nm
Mg I $b_{,1,23}$	b_2 517.3 nm
Na $D_{1,2}$	589.6, 589.0 nm
Hα	656.3 nm
Hβ	486.1 nm
CO	4.6μ

$H\alpha$, the mm continuum and the wing of Lyα.

- temperature plateau near 6 000 - 7 000 K
- sharp temperature rise beginning near 8 000 K and terminating in a thin plateau near 22 000K. From the second plateau the central portion of Lyα and the 3 cm continuum is emitted

Thus by observing in different lines or even in different depths of a particular line, one can probe the chromosphere at different height levels. As it is indicated above, it is possible to observe the chromosphere in radio waves at mm to cm wavelengths. The emission processes here are free free transitions of electrons with a Maxwellian distribution.

When analyzing the H and K lines bright grains are detected. These bright grains are produced by shocks near 1 Mm (10^6 m) height in the chromosphere.

3.3.2 Radiative Transfer in the Chromosphere

Above the temperature minimum, the spectral lines are formed under non local thermodynamic equilibrium conditions (NLTE).

Let us start with the change of the specific intensity I_ν along a short distance ds: there will occur absorption and emission, both of which are described by the coefficients:

- κ_ν absorption coefficient
- η_ν emission coefficient

For simplicity we consider a homogeneous, plane-parallel atmosphere stratified by gravity. Then, the properties depend only on the height z. The surface of the atmosphere in a strict mathematical sense is where no interactions take place, i.e. the particle densities are extremely low. The optical depth is defined by:

$$d\tau_\nu = -\kappa_\nu dz, \qquad \tau_\nu = -\int_\infty^z \kappa_\nu dz' \tag{3.19}$$

The source function is the ratio between the two coefficients:

$$S_\nu = \eta_\nu / \kappa_\nu \tag{3.20}$$

In local thermodynamic equilibrium (LTE) we have the relation:

$$S_\nu = B_\nu(T) \tag{3.21}$$

which is called Kirchhoff's law, $B_\nu(T)$ being the Planck function. We can progress to solve the transport equation:

$$I_\nu(\tau_\nu = 0, \mu) = \int_0^\infty S_\nu(\tau_\nu{}') e^{-\tau_\nu{}'/\mu} d\tau_\nu{}'/\mu \tag{3.22}$$

In this equation $\mu = \cos\theta$, θ being the angle between the normal to the disk center and the point where observations are done.

From a Taylor series expansion of S_ν about a not specified τ_ν^* one gets

$$I_\nu \sim S_\nu(\tau_\nu) = \mu \tag{3.23}$$

where τ_ν^* was specified to μ. That means, one observes under the angle θ to z approximately the source function at optical depth $\tau_\nu = \mu$.

Let us consider two energy levels in an atom which have the quantum numbers l (lower level) and u (upper level). The number of atoms per cm^3 in the lower level is N_l and in the upper level N_u. Of course a transition from l to u corresponds to an absorption process, where a photon of energy $h\nu_{l,u} = \chi_u - \chi_l$ is absorbed. Thus the number of transitions per cm^3 is given by:

$$n(l \to u) = N_l J_{\nu(l,u)} B(l,u) \tag{3.24}$$

$B(l,u)$ is the transition probability for the transition $l \to u$. On the other hand let us consider the number of spontaneous transitions from $u \to l$ which is independent on the intensity J:

$$n(u \to l) = N_u A(u,l) \tag{3.25}$$

$A(u,l)$ is the transition probability for spontaneous transitions. Generally, we do not know the average intensity $J_{\nu(l,u)}$. However, in thermodynamic equilibrium it is equal to the Planck function. In thermodynamic equilibrium there is a direct balancing between the number of transitions $u \to l$ and $l \to u$ and the ratio of the occupation numbers is governed by the Boltzmann formula:

$$\frac{N_u}{N_l} = \frac{g_u}{g_l} e^{-(\chi_u - \chi_l)/kT} \tag{3.26}$$

and

$$n(l \to u) = n(u \to l) \tag{3.27}$$

$$N_l \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} B(l,u) = N_u A(u,l) \tag{3.28}$$

where we have put the Planck function. Let us also substitute the Boltzmann formula:

$$\frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} = \frac{g_u}{g_l} e^{-(\chi_u - \chi_l)/kT} \frac{A(u,l)}{B(l,u)} \tag{3.29}$$

$$= \frac{g_u}{g_l} e^{-h\nu_{u,l}/kT} \frac{A(u,l)}{B(l,u)} \tag{3.30}$$

where g_u, g_l are the statistical weights of the states u, l. This was first found by Einstein. Besides absorption and spontaneous emission also the induced emission, transitions from $u \to l$ depending on the intensity J, has to be considered. The number of induced emissions is written as:

$$n'(u \to l) = N_u B(u,l) J_{\nu(u,l)} \tag{3.31}$$

In an induced emission process, the photons emitted have the same directions and phases as the inducing photons. Thus a detailed balancing in thermodynamic equilibrium reads as:

$$N_l J_{\nu(u,l)} B(l,u) - N_u J_{\nu(u,l)} B(u,l) = N_u A(u,l) \tag{3.32}$$

and using $J_{\nu(u,l)} = B_\nu$ and the Boltzmann formula:

$$\frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \left(B(l,u) \frac{g_l}{g_u} e^{h\nu_{u,l}/kT} - B(u,l) \right) = A(u,l) \tag{3.33}$$

$$B(u,l) g_u = B(l,u) g_l \tag{3.34}$$

$$A(u,l) = B(l,u) \frac{g_l}{g_u} \frac{2h\nu_{u,l}^3}{c^2} = B(u,l) \frac{2h\nu_{u,l}^3}{c^2} \tag{3.35}$$

These relations are called Einstein transition probabilities.

$B(u,l), B(l,u), A(u,l)$ are atomic constants. Though these relations were derived from thermodynamic equilibrium, they must always hold. Therefore, they can be used to get information for excitation conditions and the source function in case we do not have thermodynamic equilibrium.

By these calculations one can understand the typical profile of the Ca II H and K lines (see Fig. 3.7). There are two intensity minima on the blue and red side of the line center (called K_{1v}, K_{1r}), towards the line center two maxima (called K_{2v}, K_{2r}) and then at the line center there is a minimum (K_3). This indicates that the temperature increases in the chromosphere. While the source function decouples from the Planck function it reaches a minimum K_1, exhibits a small maximum K_2 and finally drops towards the line center. The profile of the well known $H\alpha$ line is simpler, there is just a pure absorption. That can be explained with the structure of the H atom.

A review about the diagnostics and dynamics of the solar chromosphere can be found in Kneer and Uexküll (1999) [168].

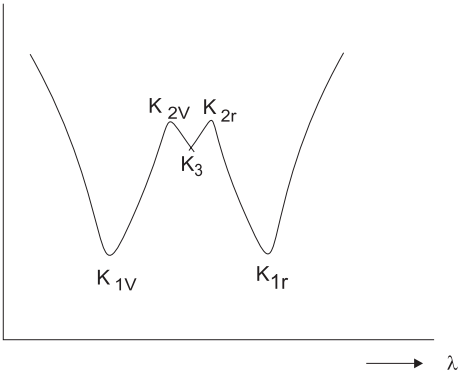


Figure 3.7: Profile of the CaII line

3.3.3 Chromospheric Heating

The temperature increases throughout the chromosphere from the temperature minimum at its base (T$\sim$ 5000 K) to several 10^4 K. The question is how does this heating mechanism work. In the reviews of Ulmschneider *et al.* (1991) [320] and Narain and Ulmschneider (1990) [229] mechanisms which have been proposed for the heating of stellar chromospheres and coronae are discussed. These consist of heating by acoustic waves, by slow and fast MHD waves, by body and surface Alfvén waves, by current or magnetic field dissipation, by microflare heating and by heating due to bulk flows and magnetic flux emergence.

Following to Kalkofen (1990) [151] the quiet solar chromosphere shows three distinct regions. Ordered according to the strength of the emission from the low and middle chromosphere they are

- the magnetic elements on the boundary of supergranulation cells,
- the bright points in the cell interior, and
- the truly quiet chromosphere, also in the cell interior.

The magnetic elements on the cell boundary are associated with intense magnetic fields and are heated by waves with very long periods, ranging from six to twelve minutes; the bright points are associated with magnetic elements of low field strength and are heated by (long-period) waves with periods near the acoustic cutoff period of three minutes; and the quiet cell interior, which is free of magnetic field, may be heated by short-period acoustic waves, with periods below one minute. This paper reviews mainly the heating of the bright points and concludes that the large-amplitude, long-period waves heating the bright points dissipate enough energy to account for their chromospheric temperature structure.

Skartlien *et al.* (2000) [287] studied the excitation of acoustic waves using three dimensional numerical simulations of the nonmagnetic solar atmosphere and

the upper convection zone. They found that transient acoustic waves in the atmosphere are excited at the top of the convective zone (the cooling layer) and immediately above in the convective overshoot zone, by small granules that undergo a rapid collapse, in the sense that upflow reverses to downflow, on a timescale shorter than the atmospheric acoustic cutoff period (3 minutes). The location of these collapsing granules is above downflows at the boundaries of mesogranules where the upward enthalpy flux is smaller than average. An extended downdraft between larger cells is formed at the site of the collapse. The waves produced are long wavelength, gravity modified acoustic waves with periods close to the 3 minute cutoff period of the solar atmosphere. The oscillation is initially horizontally localized with a size of about 1 Mm. The wave amplitude decays in time as energy is transported horizontally and vertically away from the site of the event. They also made a prediction of how to observe these "acoustic events": a darkening of intergranular lanes, which could be explained by this purely hydrodynamical process. Furthermore, the observed "internetwork bright grains" in the Ca II H and K line cores and associated shock waves in the chromosphere may also be linked to such wave transients.

The coronal heating problem can be also studied by an energy release that is associated with chromospheric magnetic reconnection. A one-dimensional circularly symmetric supergranulation reconnection model was investigated by Roald *et al.* (2000) [261] with typical quiet-Sun values. In this model, the assumed source rate of elements determines heating, because all emerged elements eventually annihilate.

As an example for observational evidence we cite the paper of Ryutova and Tarbell(2000) [267]. They analyzed spectra of CII and OVI lines corresponding to chromosphere and transition region temperatures; these showed significant broadening and complex line profiles in regions overlying the sites of small scale magnetic elements in the photospheric network. Doppler shifted multiple peaks in CII line were always seen soon after the reconnection of magnetic flux tubes occurs and usually consist of supersonic and subsonic components caused by shocks propagating upward. Multiple peaks in OVI line have more diverse features: they are not as persistent as those seen in CII line, and may have the configuration of maximum intensity peaks corresponding either to forward or reflected shocks.

Ca II H_{2V} grains can also be used as indicators for shocks. Therefore spatio-temporal correlations between enhanced magnetic fields in the quiet solar internetwork photosphere and the occurrence of Ca II H_{2V} grains in the overlying chromosphere were investigated by Lites *et al.* (1999) [197].

Cauzzi *et al.* (2000) [59] analyzed the temporal behavior of Network Bright Points (NBPs) using a set of data acquired during coordinated observations between ground-based observatories (mainly at the NSO/Sacramento Peak) and the Michelson Doppler Interferometer onboard SOHO. The NBP's were observed in the NaD_2 line and were found to be cospatial with the locations of enhanced magnetic field. The "excess" of NaD_2 intensity in NBPs, i.e. the emission over the average value of quiet regions, is directly related to the magnetic flux density. Thus in analogy with the Ca II K line, the NaD_2 line center emission can be used as a proxy for magnetic structures.

In a paper by Fossum and Carlsson it was shown that acoustic heating of the chromosphere is a factor of 10 too low to balance radiative losses (Fossum and Carlsson, 2005 [99]).

Simultaneous CaII K-line spectroheliograms and magnetic area scans were used to search for spatial correlation between the CaII K_{2V} bright points in the interior of the network and corresponding magnetic elements and 60% of the bright points spatially coincided with magnetic elements of flux density $> 4\,\mathrm{Mxcm}^{-2}$ (Sivaraman *et al.* 2000 [286].

3.3.4 Chromospheric Network, Supergranulation

On a full disk photograph taken in Ca II K a bright network surrounding darker island structures becomes visible. This pattern is known as chromospheric network. It looks like a photographic negative of the photospheric granulation pattern, however the scale is larger, typical sizes are between 20 000 and 30 000 km. This is the size of the so called supergranulation first observed by Leighton *et al.* (1962) [192]. The bright network is cospatial with the magnetic network. The supergranulation is also visible on 30 min averaged MDI Dopplergrams. Fig. 2.10 was constructed out of a full series of 7.4 hours. The frame shown is the result of averaging 30 full disk velocity maps and subtracting the contribution from the Sun's rotation. The color scale is such that dark is motion towards the observer and bright is motion away from the observer. The signature of the waves is nearly cancelled in this image since the wave periods are mostly about 5 minutes. The resulting image clearly shows the supergranulation pattern. The "smooth" area in the center is where the supergranules do not contribute to the signal since what observers see are horizontal motions and MDI measures only the component of motion directed towards or away from SOHO.

Close inspection shows that the supergranules flow outwards from their centers so that the edges towards the center are dark (motion toward SOHO) and the edges towards the Sun's limb are bright (motion away from SOHO). These flows are about 400 m/s. The typical lifetime of a supergranular cell is about half a day. Recent investigations claim a connection between boundaries of coronal holes and supergranular structures.

Random fluid motions associated with solar supergranulation may influence the interplanetary magnetic field. Magnetic footpoints anchored in the photosphere execute a random walk and the resulting magnetic variations are carried away by the expanding solar wind. The solar satellite mission Ulysses has observed the resulting large-scale magnetic-field fluctuations in the solar wind.

By spatio-temporal averaging of two-dimensional velocity measurements obtained in the MgI 5173 line November *et al.* (1981) [236] found the "mesogranulation", in order to indicate the supposed convective character of the phenomenon with a typical scale of 5 - 10 Mm and a lifetime of approximately 2 h.

The convective nature of the mesogranulation as well as the supergranulation is not sure. E.g. Rieutord *et al.* (2000) [259] assign mesogranular flows with both highly energetic granules, which give birth to strong positive divergences (SPDs) among which we find exploders, and averaging effects of data processing. A similar

explanation is suggested for the supergranulation.

Hathaway *et al.* (2000) [128] analyzed power spectra from MDI observations. The spectra show distinct peaks representing granules and supergranules but no distinct features at wavenumbers representative of mesogranules or giant cells. The observed cellular patterns and spectra are well represented by a model that includes two distinct modes - granules and supergranules.

Up to now we know that there exist three different scales of motion in the photosphere:

- Granulation: size about 1 000 km, lifetime 0.2 hr, vertical flow $\sim 1\,\mathrm{kms}^{-1}$.
- Mesogranulation: diameter 5 000 km, lifetime 3 hr, vertical flow $\sim 60\,\mathrm{ms}^{-1}$.
- Supergranulation: diameter about 32 000 km, horizontal flow $\sim 400\,\mathrm{ms}^{-1}$, lifetime 20 hr.

The scales of granulation, mesogranulation and supergranulation are discussed by Rast, 2003 [251]. It is discussed there that the downflow plume mainly describes the granular scale and that from collective advective interaction of many small scaled and short lived granular plumes the larger spatial and temporal scales of mesogranulation and supergranulation naturally arise.

3.4 Solar Flares

The first recorded observation of a flare was a local brightening in the visible light but most solar flares can be observed in the Hα line. The typical energy release is of the order of 10^{25} J within half an hour. A recent review on solar flares was given by Vrsnak, 2005 [327].

3.4.1 General Properties

Flares produce effects throughout the whole electromagnetic spectrum. They produce X rays and UV radiation which is an evidence for very high temperatures during a flare outburst. The radio waves indicate that a small fraction of the particles are accelerated to high energies. Most of the radiation is synchrotron radiation produced by electrons moving in helical paths around magnetic field lines. The flux of high energy particles and cosmic rays is also increased at the Earth as a result of an intense flare. Magnetic storms on Earth often occur with a delay of about 36 h after the flaring event was observed on the Sun. This is basically interpreted as an enhancement in the solar wind which compresses the magnetosphere and increases the magnetic field near the surface of the Earth. Flares occur in regions where there is a rapid change in the direction of the local magnetic field. The favored mechanism to explain the sudden energy release in flares is magnetic reconnection.

Let us describe the basic processes of a flare (see Vrsnak, 2005 [327]).

As shown in Fig. 3.8 two oppositely magnetic field lines interact due to a compression - reconnection occurs - and the resulting flaring loop is shown by

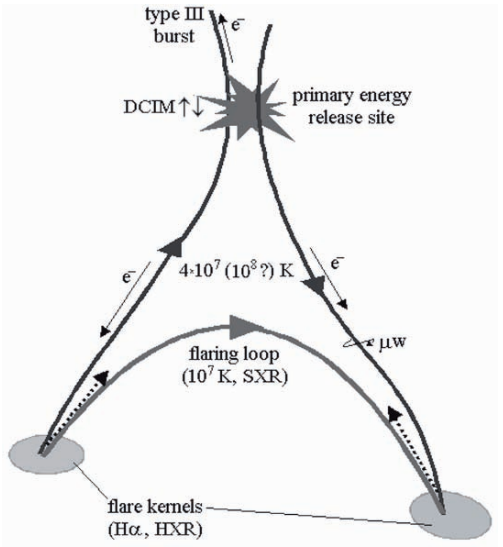


Figure 3.8: Summary of basic processes of most solar flares and their emission regions. Courtesy: B. Vrsnjak.

bold arrow-lines (grey). Electron beams are given by thin arrows (and marked by e^-). Chromospheric evaporation from the flare kernels is indicated by thick dotted arrows. As it is seen in the sketch, the primary energy release takes place in the corona at heights between 10^4 and 10^5 km. DCIM indicates fast drifting bursts in the 200-2000 MHz range. HXR (Hard X-ray) emission is related to radio features. HXR emission at successively lower energies indicates delays of slower electrons relative to faster ones. The power of flares is related to the height of the energy release site. Flares are more powerful and impulsive when the energy release site is located at low heights. This can be explained by the weakening of the magnetic field with height. As is also seen in the Fig. 3.8, electron beams that are produced at the primary energy release site can escape outwards exciting type III bursts. Electrons attached to the closed field lines become trapped between the two magnetic mirrors located near the footpoints and they excite type IV bursts.

In very strong fields also μwave emission occurs (mm-cm range). Electrons with small pitch angles penetrate through the magnetic mirrors. They hit the dense transition region and chromosphere and excite line emission of atoms and ions and hard X ray emission (HXR). This process is also called thick target Bremsstrahlung. The chromospheric plasma is heated and starts to expand. This is the evaporation process which continues until a new hydrostatic equilibrium is reached. This is a source of soft X-ray emission (SXR), the plasma has a density of $\sim 10^{-3}\,\mathrm{cm}^{-3}$ and a temperature of $\sim 10^7$ K. The evaporation and SXR emission is a cumulative effect of precipitating electrons- the cooling is relatively slow. The SXR curve behaves as a time integral of the HXR curve. Or it can also be stated that the HXR curve looks like the time derivative of the SXR curve. This is called

Table 3.3: Optical classification scheme of solar flares

Importance class	Area A at disk 10^{-6} sol. hemisphere	Energy (erg)
S	A<100	10^{28}
1	$100 \leq A < 250$	10^{29}
2	$250 \leq A < 600$	10^{30}
3	$600 \leq A < 1200$	10^{31}
4	$A \geq 1200$	10^{32}

Table 3.4: Soft x-ray classification scheme of solar flares

Soft x-ray class	Peak in power of 10 in the 0.1-0.8nm flux $\mathrm{W\,m^{-2}}$
A	-8
B	-7
C	-6
M	-5
X	-4

the Neupert effect (see also Neupert, 1968 [231] and Veronig et al., 2002 [326]).

3.4.2 Classification of Solar Flares

There are different classification schemes of solar flares:

- Optical classification: in this scheme importance classes S, 1, 2, 3, 4 are used, according to the area of the flaring region at disk center (given in millionths of a solar hemisphere, see Table 3.3). The values for the total energy released are given from [17].

  In this scheme the letter S stands for subflares.

- Soft x-ray classification: since 1970 flares are also classified based on soft x-ray observations of the Sun in the 0.1-0.8 nm band by Earth orbiting satellites. The size of the flare is given by the peak intensity (on a logarithmic scale) of the emission (see Table 3.4).

  According to Table 3.4 a B5 flare has a peak flux of $5 \times 10^{-7}\,\mathrm{Wm^{-2}}$. Flares smaller than C1 can only be detected during a solar minimum phase when the general x-ray background is low. Occasionally, flares exceed class X9 in intensity and are referred simply to as X10, X11...

- Classification into impulsive and gradual: in fully developed flares an impulsive phase is always followed by a gradual main phase. The classification according to the time scales is indicative of the magnetic topology.

Table 3.5: Radio classification scheme of solar flares

Type	Confined	Eruptive
Radio bursts	III/V	II/IV
Soft x-ray duration	< 1 h	> 1 h
CME	-	Yes
Interplanetary shock	.	Yes
Events/year	~ 1000	~ 10

Long duration flares are linked to coronal mass ejections (CMEs) but recent observations also showed that some short duration flares may have ejecta. Coronal mass ejections (CMEs) leave the Sun at speeds up to 2000 km/s and can have angular spans over several active regions whereas flares imply events that are localized within a single active region. In CMEs the magnetic field lines are opened in eruptive events. There occurs a closing down or reconnection within several hours providing a prolonged energy release that is typical for gradual or eruptive flares. The intersection of the newly formed flare loops with the solar surface can be observed: two parallel ribbons in Hα. Therefore, in the older literature we find the designation double ribbon flares for eruptive flares.

Eruptive flares are very important because of their complexity and association with geomagnetic storms.

Confined or impulsive events may also result from loop top magnetic reconnection. An impulsive flare of say 10^{24} J is typically spread over an area of several $10^{14}\,\mathrm{m}^2$ in Hα. Therefore, the main difference between eruptive and impulsive flares may be the order of intensity.

Radio bursts and flares: solar flares are associated with radio bursts which are observed at wavelengths ranging from mm to km. The radio classification scheme was developed during the 1950s by Australian and French solar radio astronomers. The different types can be easily recognized in the so called *dynamic spectrum*: in such a diagram on the x-axis the time is plotted and on the vertical axis the frequency. Since the frequency varies with height, one can easily study the evolution with height of this phenomenon that means the propagation throughout the solar corona.

The Wind spacecraft[7] observes radio bursts in the frequency range 1-14 MHz. Standard patrols of bursts are made above 25 MHz. With the Bruny Island Radio Spectrometer, this gap is filled and it is studied whether radio bursts can be used in diagnosing energetic particle generation and propagation in the inner heliosphere (Cane, Erickson, 2006 [54]).

Bursts of type III and type V are characteristic phenomena of impulsive flares (or the impulsive or initial phase of fully developed eruptive flares). Type III bursts and their associated type V continua are attributed to flare-accelerated electrons moving along open field lines into the corona. Type II and type IV bursts are most commonly identified with eruptive flares. Type IV emission is related to magnetic reconnection in CME.

[7] was launched in 1994, part of the ISTP project

Type II radio bursts result from plasma radiation associated with a MHD shock propagating through the corona (∼500 km/s). This can be observed by a slow drift emission. More than 90% of type II bursts have an associated flare. They accompany 30% of flares with an Hα importance class 2 and 3. 70% of all type II bursts are associated with a CME.

3.4.3 Where do Flares Occur?

Like all signs of solar activity, flares are associated with magnetic fields and restructuring of these fields. As a general rule, flares occur above the places in the photosphere with largest $\nabla \times B$. These are the locations where the electric current has a maximum. Preferred are regions in sunspots or groups of sunspots where new and oppositely directed magnetic flux emerges from below. Large gradual flares often occur above the neutral lines in the photosphere which separates regions with opposite magnetic polarity. Neutral lines are bridged by arcades of loops and in Hα one sees two bright ribbons formed by the footpoints on each side of the neutral line. Flares then occur above the part of the neutral line which has experienced most shear by different surface motions on both sides. In quiet regions, the most powerful microflares occur at the boundary of supergranular cells. The frozen-in magnetic field lines are swept to the down-draft region near the supergranular boundary forming the magnetic network. At time scales of a few tens of minutes these magnetic elements can be observed to appear and disappear.

Current helicity: substantial changes of current helicity distribution in an area or in its vicinity probably lead to flare eruptions. The total current helicity is defined by

$$H_c = \mathbf{B}.\nabla \times \mathbf{B} \tag{3.36}$$

A measure for the z component can be obtained from

$$h_c = \mu_0 B_z J_z \qquad J_z = \frac{1}{\mu_0}\left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right) \tag{3.37}$$

Gaizauskas (1989) [106] made a categorization of flare precursors. According to him, a precursor is a transient event preceding the impulsive phase. We give a short list here:

- Homologous flares: these are earlier flares in the same location with similar emission patterns. They occur most often in periods of frequent flare activity. The rate of repetition ranges from a few per hour to several days.

- Sympathetic flares: these group consists of earlier flares in different locations but erupting in near synchronism. From soft x-ray images of the solar corona it is evident that there exist links between even remote active regions. Studies have shown that one flare can trigger another.

- Soft x-ray precursors: these are transient enhancements in soft x-rays lasting for several minutes; they occur in loops or unresolved kernels or close to flare sites. Weak soft x-ray bursts are often observed at the time of the onset of a

CME. Sometimes several tens of minutes prior to the impulsive phase. The location is at one foot of a large coronal arch which already exists. The process can be interpreted by a small magnetic structure which interacts with the large coronal arch at one of its footpoints. The whole structure becomes then destabilized.

- Radio precursors: often tens of minutes before the onset of a flare, changes in intensity and polarity in microwaves are observed. However the correlation with flares is not very strict.

- UV precursors: small scale transient brightenings above active regions, some bright UV kernels coincide with the later flares, others do not.

- surging arches: a surging arch is a transient absorbing feature visible at wavelengths displaced from the central core of Hα. Simultaneous red- and blueshifted components are also visible. The arch is initially straight, expands and unravels in multiple strands by the time the associated flare erupts. However the link to flares is not very strong.

- Prominence eruptions: very often they precede two ribbon flares. The time delay between the onset of the prominence eruption and the impulsive phase is of the order of minutes. Enhanced mass motion, a slow rise of the prominence and untwisting can precede the main flare by hours.

Of course in all the cases joint observations covering the whole electromagnetic spectrum are important. In the review given by Aschwanden *et al.* (2001) [14] the authors focussed on new observational capabilities (Yohkoh, SoHO, TRACE).

The formation of a radio-emitting shock wave and its precursor above a flaring active region was investigated in Klassen *et al.* (2003)[166]. They used imaging and spectral observations of radio bursts with Yohkoh soft hard and X-ray imaging observations and identified type II precursor as a signature of the reconnection process above the expanding soft X-ray loops.

Characteristics of flare producing sunspot groups were discussed by Ishii *et al.* (2000) [146]. A review about reconnection theory and MHD of solar flares is given by Priest (2000) [249].

3.4.4 Prominences

Prominences are great areas of luminous material extending outwards from the solar atmosphere and were first observed during eclipses. They can also be observed in the light of Hα. Over the photosphere they appear as dark filaments, at the limb as bright structures.

The prominence plasma contains 90% of hydrogen which is partially ionized in the central coolest parts of prominences where the temperatures are between 6000 and 8500 K or maybe even lower. At the boundary of prominences the temperature rapidly increases to coronal values (more than 1 million K). The plasma density in the central cool parts is about two orders of magnitude larger than that in the corona.

These facts imply that the magnetic field is crucial for the prominence support and stability. The intensity of the field ranges up to a few tens of Γ.

Some prominences are short lived eruptive events (variations within minutes to hours), others can be quiescent and survive many rotational periods of the Sun. The upper parts are often located in the hot corona. Quiescent prominences appear as huge arches of dense cool material embedded in the hot corona. The length of the arch is typically several 100 000 km and the height up to 10^5 km. A quiescent prominence may change into an eruptive prominence. The typical thickness of the loop is 10^4 km. At the end of its life, a prominence disperses and breaks up quietly or it becomes eruptive or matter falls back down the field lines to the photosphere. The particle densities range from $10^{16...17}\,\mathrm{m}^{-3}$ which is a hundred times greater than coronal values.

Prominences are mostly located along the so-called neutral lines where the vertical photospheric magnetic field changes its sign. Along the neutral line, the vertical component of that field is zero.

A possible mechanism to understand cool prominence material (temperature about $10^4\,\mathrm{K}$) is *thermal instability*. The equilibrium of the corona requires:

$$\text{heating} = \text{cooling} \tag{3.38}$$

Suppose now that this equilibrium is disturbed locally. The density of the corona increases in such a disturbed region and it will become cooler than its surroundings. If we assume that thermal conduction from the hotter surroundings cannot restore equality of temperature, the dense region will continue to cool until it reaches a new equilibrium in which its heat input balances its heat output. When a magnetic field is present, particles can only move along the field lines, this means that thermal conductivity parallel to the field lines is very much greater than $\kappa_\perp$. As a result, the longest dimension of any cool material is likely to be along the field. The equation of equilibrium of a magnetized fluid acted on by a gravitational field, g, in the z-direction is:

$$\mathbf{0} = -\mathrm{grad}\mathbf{P} - \rho\mathbf{g}\bar{\mathbf{z}} - \mathrm{grad}(\mathbf{B^2}/\mathbf{2}\mu_\mathbf{0}) + \mathbf{B}.\nabla\mathbf{B}/\mu_\mathbf{0} \tag{3.39}$$

The perfect gas law:

$$P = \Re\rho T/\mu \tag{3.40}$$

where $\Re$ is the gas constant and μ the molecular weight. In a simple model Kippenhahn and Schlüter (1957) [163] assumed that the temperature T and the horizontal magnetic field components $B_\mathrm{x}, B_\mathrm{y}$ were constant and that P, ρ and B_z were functions of x alone. The prominence is represented as a plane sheet.

Tripathi *et al.*, 2004 [316] studied an erupting prominence with EIT and then with LASCO when it developed into a CME.

A recent review about prominences was given by Heinzel and Anzer, 2005 [131].

Table 3.6: Tomography of the solar corona by observations at different radio frequencies

ν MHz	λ cm	$F_\odot$ 10^{-22} Wm^{-2}Hz^{-1}	T
30	1000	0.17	5.1×10^5
300	100	14.9	7.0×10^5
3000	10	69	31 000
30 000	1.0	1862	10000
300 000	0.1	113 200	5900

3.5 The Corona

3.5.1 Basic Facts

During a total solar eclipse, when the moon occults the Sun for a few minutes we can observe the outer atmospheric layers of the Sun, the chromosphere and the corona the latter extending far out. There are possibilities to observe the corona when there is no total eclipse. With a coronagraph the light from the photosphere is occulted and blocked out by a disk placed inside the telescope. Space observations allow a continuous monitoring of the corona in the UV and EUV. The shape of the corona which extends to several solar radii depends on the sunspot cycle being more spherical around the Sun at solar maximum.

The corona includes open streamers and closed loops. These phenomena are associated with magnetic field lines. Those which return to the surface of the Sun provide closed loops, the open streamers are related to field lines which extend to a large distance from the Sun carrying the solar wind, which is a continuous mass loss of the Sun. The light from the solar corona was very puzzling since many strong spectral lines could no be identified when discovered (such as Helium or Coronium; therefore their names). Later it was clarified that many of these lines are forbidden lines arising from a transition in which an electron can spend an unusually long time in an excited state before it returns to the ground level. Under normal laboratory conditions the atom will undergo many collisions and the electron will either move to the ground state without emission or move to a higher level. Therefore, no forbidden lines will be observed. In the corona the density of matter is extremely low, collisions are infrequent and forbidden transitions can be observed[8].

Moreover, the coronal spectrum contains lines from highly ionized atoms indicating kinetic temperatures of several 10^6 K which was a big surprise when discovered. Typical lines are Ca XII... Ca XV, Fe XI...Fe XV etc. Here the roman numeral is one more than the number of electrons removed from the atom. E.g. Ni XVI has lost 15 of its 28 electrons.

In Table 3.6 it is demonstrated that the corona can be observed by radio emission in different wavelengths. The lower the wavelength, the deeper the zone

[8]This is also well known for gaseous nebulae in astrophysics

where the emission occurs, thus also the deeper the temperature. The values are given for the quiet Sun (Landolt, 1981 [17]).

3.5.2 Observational Features in the Corona

The most important features seen in the corona are:

- Coronal loops are found around sunspots and in active regions in the corona. These structures are associated with the closed magnetic field lines that connect magnetic regions on the solar surface. As it is shown in the chapter on MHD, in the corona the magnetic field dominates the motion of the plasma, and therefore the plasma is aligned in magnetic loops. These loops last for days or weeks. Some loops, however, are associated with solar flares and are visible for much shorter periods. These loops contain denser material than their surroundings. The three-dimensional structure and the dynamics of these loops is investigated for that reason.

- Helmet streamers are large cap-like coronal structures with long pointed peaks. They are found usually over sunspots and active regions. Often a prominence or filament lying at the base of these structures can be seen. Helmet streamers are formed by a network of magnetic loops that connect the sunspots in active regions and help suspend the prominence material above the solar surface. The closed magnetic field lines trap the electrically charged coronal gases to form these relatively dense structures. The pointed peaks are formed by the action of the solar wind blowing away from the Sun in the spaces between the streamers.

- Polar plumes are long thin streamers that project outward from the Sun's north and south poles. At the footpoints of these features there are bright areas that are associated with small magnetic regions on the solar surface. These structures are associated with the "open" magnetic field lines at the Sun's poles. The plumes are formed by the action of the solar wind in much the same way as the peaks on the helmet streamers.

- Coronal Holes: From X-ray observations it was seen that the temperature of the corona is not uniform. The lower temperature regions are called coronal holes. They are particularly prominent near sunspot minimum and near the solar poles. Coronal holes tend to form near the centers of large unipolar magnetic regions; a comparison of the X-ray images with those of magnetic field lines calculated on the assumption that the observed photospheric field line structures extend into the corona as potential fields, indicates that they are regions of open (diverging) magnetic fields. Coronal holes can also be observed in spectroheliograms taken in the 1083.0 nm line of Helium. They tend to rotate more slowly than sunspots or supergranular patterns and not differentially.

  The fast-speed solar wind originates form the coronal holes (e.g., Krieger *et al.*, 1973) [175], and accordingly they are considered the main reason for the

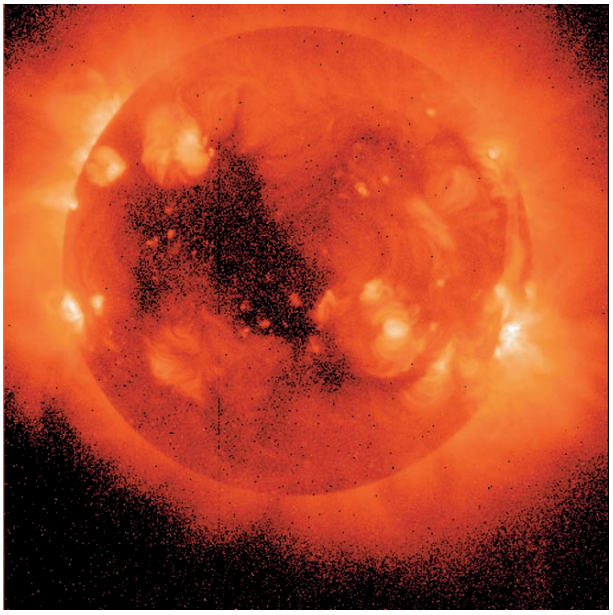


Figure 3.9: Coronal hole seen by the solar satellite YOHKOH

"recurrent" type of geomagnetic activity. They may form at any latitude. For the solar cycle of greatest importance are the unipolar coronal fields. When the polar fields are strongest during sunspot minimum polar coronal holes are well defined. They disappear during the polar field reversals near sunspot maximum.

3.5.3 Coronal Mass Ejections, CME

Coronal mass ejections, CMEs are the most energetic events in the solar system. Coronal material of mass up to 10^{16} g is expelled at speeds of several 10^2 to 10^3 kms^{-1} from the Sun. CME like structures have been seen in historical eclipse drawings. But it was recognized from space born coronograph observations like OSO-7 and Skylab, that these are features that are expelled from the corona. First they were called coronal transients. In 1976 the term Coronal Mass Ejection appeared. Gosling *et al.* 1976 [117] observed 66 such events during the Skylab mission (May 1973-Jan 1974). They also noticed that the speeds of these events (they found values between 100 and 1000 km/s) seem to be somehow correlated with the activity in Hα and statistics indicate that the fastest mass-ejection events tend to produce type II-IV burst pairs, while single type II or type IV bursts tend to be associated with events of intermediate speed. They also report on clouds observed at a distance of 1 AU[9]which seem to be related to CMEs.

[9]1 AU = 1 Astronomical Unit, mean distance Earth-Sun, 150×10^6 km

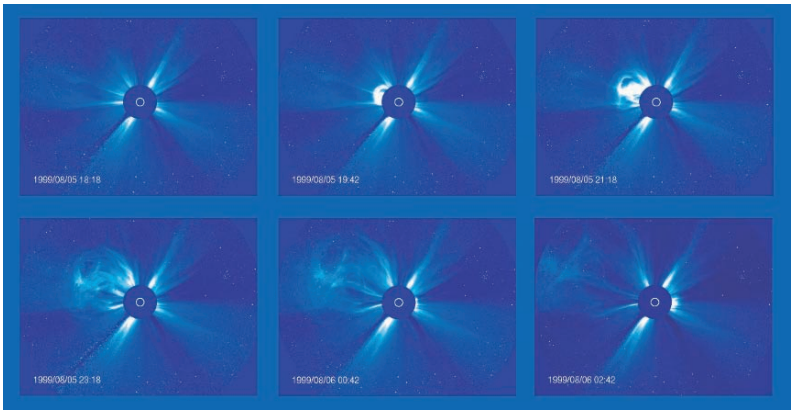


Figure 3.10: Progress of a Coronal Mass Ejection (CME) observed over an eight hour period on 5-6 August 1999 by LASCO C3. The dark disk blocks the Sun so that the LASCO instrument can observe the structures of the corona in visible light. The white circle represents the size and position of the Sun. Courtesy: SOHO/LASCO. SOHO is an ESA/NASA mission.

CMEs very often appear in a three part structure:

- bright frontal part,
- darker cavity or void,
- the core, frequently the brightest structure.

Such a structuring is seen best when CMEs erupt close to the solar limb- then they are seen from the side. Earth- (or oppositely) directed CMEs show an outflow and expanding brightness around the Sun- these are called halo CMEs (see also Jackson *et al.*, 2002 [147], where it is discussed whether Halo CMEs will hit or miss the Earth). Therefore, Halo CMEs are of special interest for space weather.

CMEs can be observed in white light[10]. In white light we see photospheric light scattered on coronal free electrons (Thomson scattering). The brighter the structure, the more massive. Brightness does not mean temperature. They can also be observed in other wavelengths, where near surface structures are observed (Hα, He 1083 nm, EUV, X-rays, microwaves to radio).

How often do CMEs occur? SOHO observations[11] yield the following frequencies of CME occurrences:

- solar activity minimum: $\sim$0.5 day^{-1}
- solar activity maximum: $\sim$ 4.5 day^{-1}.

[10]Note that the photospheric light is 10^6 times brighter than the corona
[11]see also the SOHO/LASCO catalogue: http://cdaw.gsfc.nasa.gov/CME_list/

The CME mass shows no cycle dependence, whereas the cycle influences their latitudinal distribution: during minimum CMEs are concentrated around the equator, during maximum they originate from a wide range of latitudes.

There exist two types of CMEs:

1. flare related CMEs

2. CMEs associated with filament eruption.

Flare associated CMEs are, on average, faster (median speed 760 km/s) than the ones associated with filament eruption without flare (median speed 510 km/s). The temperature is about 8000 K in the core and more than 2 million K in the front part and in the cavity.

SOHO/LASCO data from 1996 to 2001 were collected by Yeh *et al.*, 2005 [343] and they showed that the observed CMEs reveal a similar power-law behavior as flares, and the power-law indices for both phenomena are almost identical. This finding strongly supports the viewpoint that solar flares and CMEs are different manifestations of the same physical process.

CMEs are an important factor in coronal and interplanetary dynamics by injecting large amounts of mass and magnetic fields into the heliosphere causing geomagnetic and interplanetary shocks which is a source of solar energetic particles.

The geoeffectiveness of CMEs is reviewed in the paper by Webb, 2002 [332] and Kim *et al.*, 2005 [162]. They considered more than 7000 CMEs observed by SOHO/LASCA and also 300 frontside halo CMEs between 1997 and 2003. The geomagnetic storm that is associated with the CME was measured by the Dst index (see next chapter). They found that the probability of front side CME geoeffectiveness is 40%. For speeds >400 km/s and $L < 50^0$ the probability of detection is high (80%) but also the false alarm rate is high (60%). The most probable areas (or coverage combinations) whose geoeffectiveness fraction is larger than the mean probability (about 40%), are $0^0 < L < +30^0$ for slower speed CMEs ($\leq 800\,\mathrm{km/s}$), and $-30^0 < L < +60^0$ for faster CMEs (>800 km/s). Manchester *et al.*, 2004 [205] gave a study of a numerical simulation of a CME propagating form the Sun to 1 AU. They found that CME is very effective in generating strong geomagnetic activity on Earth through a strong sustained southward Bz and by a pressure increase associated with the CME driven shock that compresses the magnetosphere.

A recent review about CMEs can be found in van Driel-Gesztelyi, 2005 [324].

3.5.4 Heating of the Corona

As it has been described already, the temperature increases from the solar surface (photosphere 6000 K) to the corona (several 10^6 K). Therefore, there must be some heating process responsible for that.

Two basic facts of the corona thus have to be taken into account if we want to explain its heating:

- hot temperature
- low density (only about 10^{-12} that of the photosphere).

The original idea for the heating of the corona was entirely non-magnetic. From laboratory experiments we know that if a fluid is set into violent motion, it emits sound with the amount of sound rising as a high power of the average velocity of the fluids. As we have seen, in the photosphere convective motions occur. If these convective motions produce sound waves, they must propagate outwards from the surface of the Sun. The wave motion has an energy density of

$$E_{\text{wave}} = \frac{1}{2}\rho v^2 \tag{3.41}$$

This energy is conserved. If the wave moves into a region of lower density, then the wave amplitude must increase. The wave turns into a shock wave and there is a strong dissipation of energy. This is converted into heat and the local temperature increases.

However, it turned out that a purely acoustic heating of the corona is not sufficient to explain the high temperatures there. Acoustic heating may be important in the outer layers of some stars.

Today[12] we assume that the following two processes are the main reason for the hot corona:

- MHD waves: as it has been outlined, when a magnetic field is present there are two characteristic speeds of wave propagation, the sound speed c_{s} and the Alfvén speed c_{H}. If $c_{\text{s}} >> c_{\text{H}}$ magnetic effects are negligible but this is not the case for the outer solar atmosphere. The heating process by MHD waves is analogous to the above mentioned acoustic heating. But it has to be stressed that MHD waves have an anisotropic propagation.
- Magnetic reconnection: The footpoints of magnetic fields often are seen to be anchored in the photosphere. In this region they are being continually moved around by convective motions. Thus magnetic reconnection occurs and electric currents flow which are dissipated.

There seems to be two problems with that interpretation. MHD waves cannot carry enough energy through the chromosphere to the corona and Alfvén waves dissipate their energy very fast when entering the corona. Bogdan *et al.*, 2003 [41] have shown that Alfvén waves can transmute into other wave modes at the base of the corona.

The first observational evidence of the presence of waves in the corona was made by SOHO EUV observations. Waves with a frequency of 1 mHz were found but they could only contribute to about 10% of the needed energy. The photosphere is covered by small magnetic elements (size below 1000 km). These small elements are constantly perturbed by granulation motions. The magnetic field in the corona that is anchored at these elements therefore constantly is perturbed

[12] see also the book: Mechanisms of Chromospheric and Cornal Heating, P. Ulmschneider, E. Priest, R. Rosner, 1991, Springer

and reconnection occurs due to the motion of the magnetic carpet. Maybe a series of mircoflares occurs (see e.g. Benz, 2003 [33]).

RHESSI studied gamma and X-ray emission from flares and mircoflares. Microflares emit hard X-ray and it turned out that mircoflares are quite similar to normal flares. The type III radio bursts seem to be in relation with series of microflares- the radio signals decrease in frequency like the whistle from a departing train. In type III bursts electrons are accelerated in open magnetic field lines and the particles escape from the Sun. RHESSI observations of microflares were found to be coincident with TRACE observations (showing jets in the EUV) (see Liu *et al.* 2004 [198]).

3.6 Solar Wind and Interplanetary Magnetic field

3.6.1 Diagnostics of the Solar Wind

The Sun loses continuously mass and this mass loss is called solar wind[13]. The existence of the solar wind was first suggested to understand magnetic storms on the Earth. During magnetic storms, the properties of the Earth's ionosphere are modified and radio communication can seriously become disrupted for some time (about 36 hours) after the observation of some violent activity on the Sun (flare). Such a perturbation cannot be caused by electromagnetic radiation from the Sun because it takes 8 minutes to reach the Earth. Therefore, it was suggested that the Sun was emitting particles which caused magnetic storms when they reach the neighborhood of the Earth.

In that context it is interesting to remark that it was Carrington who discovered in September 1859 a white light flare and then 4 hours after midnight there commenced a great magnetic storm on the magnetic instruments[14].

Another hint for the existence of a solar wind arose from observations of comet tails (this was first studied by Biermann in the 1950). These are produced when comets are close enough to the Sun and the tails always point away from the Sun. Originally, it was believed that radiation pressure produces the tails. If small particles in the comet absorb radiation from the Sun they take up energy and momentum. If they subsequently emit radiation, this emission is isotropic into all directions and this will carry off no momentum- the matter will be pushed away from the Sun and thus the dust tails are produced. But observations showed that there is also a plasma tail consisting of ionized gas. If the Sun emits a continuous stream of plasma, the ionized solar gas would collide with atoms - momentum is transferred and charge exchange reaction occur: an electron will be exchanged between an incoming charged particles and a neutral cometary particle which produced the plasma tail. Since the charged particles move around magnetic field lines, the plasma tail is aligned with the local interplanetary field.

[13] see e.g. the classical textbook: A. J. Hundhausen, Coronal Expansion and Solar Wind, 1972, Springer

[14] He reported this observation to the Royal Astronomical Society

Figure 3.11: Comet Hale Bopp (1997); the fainter ion tail is clearly seen.

Satellite Measurements

The first in situ measurement of the solar wind was made in 1962 by Mariner 2[15]

First we want to mention that besides SOHO two satellite missions measure the solar wind: Ulysses and ACE. Ulysses was launched from the space shuttle Discovery in 1990. The spacecraft made a journey to Jupiter where the giant planet's gravity pulled the spacecraft into a trajectory that carried it over the Sun's south pole in the fall of 1994 and its north pole in the summer of 1995. The next passes over the Sun's south pole occurred during 2000 and over the north pole during 2001. These two orbital passes provide views of the solar wind at times near the minimum of solar activity and the maximum of solar activity. It was found that in 2000 the south magnetic pole almost completely vanished at the time of solar maximum. In November 2006 Ulysses will continue with a third south polar pass and beginning of December 2007 with its third north polar pass.

The solar wind speed, magnetic field strength and direction and composition were measured.

The Advanced Composition Explorer (ACE) satellite was launched in August of 1997 and placed into an orbit about the Lagrangian L_1 point between the Earth and the Sun[16]. ACE has a number of instruments that monitor the solar wind.

The SOHO/SWAN experiment (Solar Wind ANisotropies) measures the Lα radiation that is scattered by hydrogen atoms, which flow into the solar system. This scattered radiation is called interplanetary Lyman alpha radiation and SWAN observes interplanetary Lyman alpha radiation from all directions of the sky. These Hydrogen atoms collide with solar wind protons and get ionized. This yields to an ionization cavity around the Sun. But the form and shape of this cavity is dependent on the solar wind. Therefore the measurement of the interplanetary

[15]Mariner 2 also detected the slow retrograde rotation rate of Venus, its surface temperatures and high surface pressures

[16]The L_1 point is one of several points in space where the gravitational attraction of the Sun and Earth are equal and opposite located about 1.5 million km from the Earth in the direction of the Sun

UV Lα glow permits to determine the solar wind latitudinal distribution. If the solar wind were isotropic, the hydrogen distribution and the Lyman alpha emission pattern would be axisymmetric around the direction where the interplanetary hydrogen flows into the solar system. However, this is not true.

The chemical composition of the solar wind is interesting to investigate since it gives us hints about its origin, i.e. the sources. The most important fact is that the solar wind composition is different from the composition of the solar surface and shows variations that are associated with solar activity and solar features (Bochsler, 2001 [39]).

Also *magnetic clouds* have been observed in the solar wind. These are produced when solar eruptions (flares and coronal mass ejections) carry material off of the Sun along with embedded magnetic fields. These magnetic clouds can be detected in the solar wind through observations of the solar wind characteristics - wind speed, density, and magnetic field strength and direction.

References on magnetic clouds can be found in Burlaga *et al.* (1981) [49]. About one half of all magnetic clouds have (and usually drive) upstream interplanetary shocks, or steep pressure pulses, that in most cases possess large energy- and dynamic pressure-increases across their ramps in a stationary frame of reference. When such a sharp upstream pressure increase encounters the Earth's magnetosphere it pushes it in causing a major reconfiguration of its boundary current system measured on the ground usually some (5-10) hours before the start of the main phase of a magnetic storm (Lepping, 2001 [194]).

Planetary Magnetospheres

The Earth's magnetosphere will be described in detail in subsequent chapter.

Here we briefly outline measurements of the magnetic fields of other planets which are useful as diagnostics of the solar wind[17]. The magnetic field of Mercury and the structure and dynamics of Mercury's magnetosphere are strongly influenced by the interaction of the solar wind with Mercury. In order to understand the internal magnetic field, it will be necessary to correct the observations of the external field for the distortions produced by the solar wind. The satellites Helios 1 and 2 made a number of passes in the region traversed by the orbit of Mercury; thus it was possible to investigate the solar wind environment of Mercury. The variables that govern the structure and dynamics of the magnetospheres of Mercury and Earth are approximately 5-10 times larger at Mercury than at Earth. Thus, the solar wind interaction with Mercury will be much stronger than the interaction with Earth (Burlaga, 2001 [50]). The solar wind is not constant and since Mercury is closer to the origin of it, the solar wind at Mercury is probably more variable than that at Earth.

Mercury, Earth, Jupiter, Saturn, Uranus, Neptune, and Ganymede (satellite of Jupiter), have presently-active internal dynamos while Venus, Mars, at least two of the Galilean moons, the Earth's moon, comets and asteroids do not. These active dynamos produce magnetic fields that have sufficient strength to stand off

[17]See also e.g.: Solar Wind- Magnetospheric Coupling, Y. Kamide, 1986, Kluwer

the pressure of the exterior plasma environment and on the other hand interesting interactions with the solar wind can be studied. Moreover, e.g. the jovian magnetosphere includes a strong time-varying energy source that adds to the dynamics of its magnetosphere and produces a quite different circulation pattern than that found at Earth and, presumably, Mercury. Also the non magnetized planets Venus, Mars and even comets have induced magnetospheres associated with the solar wind interaction with their atmospheres. Cometary magnetospheres, parts of which can be remotely sensed, exhibit spectacular disruptions called tail disconnections. Even the atmosphereless bodies with weak magnetic fields can interact with the solar wind. Small magnetic anomalies on the moon and possibly asteroids cause weak deflections of the solar wind. This is discussed in the paper of Russell (2001) [265].

Krymskii *et al.* (2000) [176] investigate the interaction of the interplanetary magnetic field and the solar wind with Mars. Data from the Mars Global Surveyor mission have shown that localized crustal paleomagnetic anomalies are a common feature of the Southern Hemisphere of Mars. The magnetometer measured small-scale magnetic fields associated with many individual magnetic anomalies (magnitudes ranging from hundreds to thousands nT at altitude above 120 km). Thus Mars is globally different from both Venus and Earth. The data collected by Lunar Prospector near the Moon were interpreted as evidence that above regions of inferred strong surface magnetic fields on the Moon the solar wind flow is deflected, and a small-scale mini-magnetosphere exists under some circumstances. With a factor of 100 stronger magnetic fields at Mars and a lower solar wind dynamic pressure (because of the greater distance), those conditions offer the opportunity for a larger size of small 'magnetospheres' which can be formed by the crustal magnetic fields. The Martian ionosphere is controlled both by solar wind interaction and by the crustal magnetic field. Therefore, the nature of the Martian ionosphere is probably different from any other planetary ionospheres, and is likely to be most complicated among the planetary ionospheres (Shinagawa, 2000 [283]).

3.6.2 Solar Wind and Interplanetary Magnetic Fields

The global solar wind structure from solar minimum to solar maximum is reviewed by Gibson (2001) [109].

E.N. Parker predicted the existence of a solar wind from theoretical arguments showing that a hot corona would imply a continuous stream of plasma.

There are several types of solar wind (see Table 3.7)

The solar wind varies in strength through the solar activity cycle. It has an average speed at the Earth of about 400 km/s. The total mass loss is a few $10^{-14}\,M_\odot$/yr. This is about 1 million tons of solar material flung out into space every second. If the solar wind was the same in the past then today the total mass loss of the Sun over that period would be in the order of $10^{-4}\,M_\odot$[18].

The solar wind flows along the open (the term open magnetic field lines does not imply magnetic monopoles but means that they are closed very far from the Sun in the interplanetary space) magnetic field lines which pass through coronal

[18]This mass loss rate is comparable with that due to nuclear reactions

Table 3.7: Several types of solar wind.

Component	velocity km/s	density $10^{-6}\,m^{-3}$	He %	remarks
fast	400-800	3	3-4	coronal holes quiet Sun
slow				
minimum	250-400	11	<2	often at sector boundary
maximum	250-400	11	~ 4	turbulent, shock waves
CMEs	400-2000		He^{++} up to 30	shock waves

holes. Additionally to the solar wind, the Sun also looses mass by *coronal mass ejections* (CME's). Some of them but not all are accompanied by solar flares. Low speed winds come from the regions above helmet streamers we have discussed above while high speed winds come from coronal holes. However, if a slow moving stream is followed by a fast moving stream the faster moving material will interact with it. This interaction produces shock waves that can accelerate particles to very high speeds.

As the Sun rotates these various streams rotate as well (co-rotation) and produce a pattern in the solar wind much like that of a rotating lawn sprinkler. At the orbit of the Earth, one astronomical unit (AU) or about 1.5×10^8 km from the Sun, the interplanetary magnetic field makes an angle of about 45 degrees to the radial direction. Further out[19]the field is nearly transverse (i.e. about 90 degrees) to the radial direction.

The Sun's magnetic field, that is carried out into interplanetary space is called the interplanetary magnetic field, IMF. The interplanetary field lines are frozen in the plasma. Because of the Sun's rotation, the IMF like the solar wind travels out in a spiral pattern. This can be compared to the pattern of water sprayed from a rotating lawn sprinkler. The winding up of the magnetic field is named Parker spiral after the scientist who first described this (see Fig.3.12). Furthermore, sectors (typically four) with alternating inward and outward directed magnetic fields can be identified.

As the solar wind expands, its density decreases as the inverse of the square of its distance from the Sun. At some large enough distance from the Sun (in a region known as the *heliopause*), the solar wind can no longer "push back" the fields and particles of the local interstellar medium and the solar wind slows down from 400 km/s to perhaps 20 km/s. The location of this transition region (called the heliospheric termination shock) is unknown at the present time, but from direct spacecraft measurements must be at more than 50 AU. In 1993 observations of 3 kHz radiation from Voyagers 1 and 2 have been interpreted as coming from a radio burst at the termination shock. This burst is thought to have been triggered by an event in the solar wind observed by Voyager 2. From the time delay between this triggering event and the observation of the 3 kHz radiation, the distance of the termination shock has been put between 130 and 170 AU.

[19]at the orbit of Saturn

As it has been stated already, the particle density of the solar wind varies. From May 10-12, 1999, the solar wind dropped to 2% of its normal density and to half of its normal speed. This severe change in the solar wind also changed the shape of Earth's magnetic field and produced an unusual auroral display at the North Pole.

Let us give some theoretical arguments of the solar wind and describe its properties in more detail. Suppose the hot corona sits in static equilibrium on the top of the solar atmosphere. In such a case the pressure gradient in the corona must be balanced by the gravitational attraction of the Sun:

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{GM_\odot}{r^2} \tag{3.42}$$

In this equation we have replaced the variable M by $M_\odot$ since the mass of the corona is negligible to the total mass of the Sun. We also can write:

$$P = nkT_{\mathrm{kin}} \qquad \rho = nm \tag{3.43}$$

n is the number of particles per unit volume and m is the average particle mass. Please also note that T_{kin} is the kinetic temperature of the corona which is far from thermodynamic equilibrium.

In the corona, conduction is important for energy transport and if κ is the coefficient of heat conduction, then

$$\kappa = \kappa_0 T_{\mathrm{kin}}^{5/2} \tag{3.44}$$

where κ_0 is constant. If there is no inertial release of heat in the corona, the outward flow of heat L_{cond} must be constant:

$$L_{\mathrm{cond}} = -4\pi r^2 \kappa_0 T_{\mathrm{kin}}^{5/2} \mathrm{d}T_{\mathrm{kin}}/\mathrm{d}r = \mathrm{const} \tag{3.45}$$

This equation can be integrated:

$$T_{\mathrm{kin}}/T_{\mathrm{c}} = (r_{\mathrm{c}}/r)^{2/7} \tag{3.46}$$

where $r_{\mathrm{c}}, T_{\mathrm{c}}$ are radius and temperature at some point in the corona. Combining all four above equations one gets P and n as a function of r. When expanding this to the Earth one gets a kinetic temperature of $5 \times 10^5\,\mathrm{K}$ and a particle density of $4 \times 10^8\,\mathrm{m}^{-3}$. Parker pointed out that a solution of such a system to the edge of the solar system gives nonsense. At large values of r the value of P becomes constant, so that $\rho \sim r^{2/7}$. This is higher than the pressure of the interstellar medium and thus a static model of the corona does not make sense.

If the material of the corona moves outward with a velocity v_{r} in the radial direction, then equation 3.42 together with 3.43 becomes

$$nmv_{\mathrm{r}} \frac{\mathrm{d}v_{\mathrm{r}}}{\mathrm{d}r} = \frac{\mathrm{d}}{\mathrm{d}r}(nkT_{\mathrm{kin}}) - \frac{GnmM_\odot}{r^2} \tag{3.47}$$

Mass conservation requires:

$$nr^2 v_{\mathrm{r}} = \mathrm{const} \tag{3.48}$$

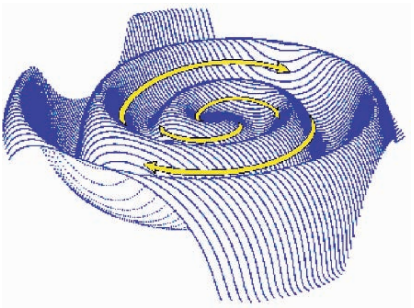


Figure 3.12: The Parker spiral. Courtesy: J. Jokipii, University of Arizona

3.6.3 High Speed Solar Wind

The high speed solar wind emanating from large coronal holes requires additional energy. It has been shown that Alfvén waves from the Sun can accelerate the solar wind to these high speeds. The Alfvén speed in the corona is quite large and therefore Alfvén waves can carry a significant energy flux even for a small wave energy density. These waves can therefore propagate through the corona and inner solar wind. The wave velocity amplitude in the inner corona must be 20-30 km/s. In the corona and inner solar wind region, the flow speed is much smaller than the Alfvén speed and the solar wind flow and the wave energy transport are along the magnetic field lines. In this region, the wave energy flux F in a magnetic flux tube is approximately constant:

$$F = \rho \Delta v^2 v_A A \tag{3.49}$$

ρ... mass density, $\sqrt{\Delta v^2}$ wave velocity amplitude, v_A Alfvén speed and A is the cross section of the flow tube. The magnetic flux $\Phi = BA$ is constant, so that the wave velocity amplitude changes with density as

$$\Delta v^2 = \Delta v_0^2 \sqrt{\rho/\rho_0} \tag{3.50}$$

The subscript 0 indicates a reference level in the inner corona.

3.6.4 Heliospheric Current Sheet

Along the plane of the Sun's magnetic equator, the oppositely directed open field lines run parallel to each other and are separated by a thin current sheet known as the "interplanetary current sheet" or "heliospheric current sheet" (see Fig. 3.13). The rotational axis and the magnetic axis of the Sun do not coincide. Therefore the current sheet becomes tilted and shows a wavy ("ballerina skirt")-like structure as it extends into interplanetary space. Therefore, the Earth is located sometimes below and sometimes above that rotating current sheet and it experiences periodic changes in the polarity of the IMF. These periods of alternating positive and negative polarity are known as magnetic sectors.

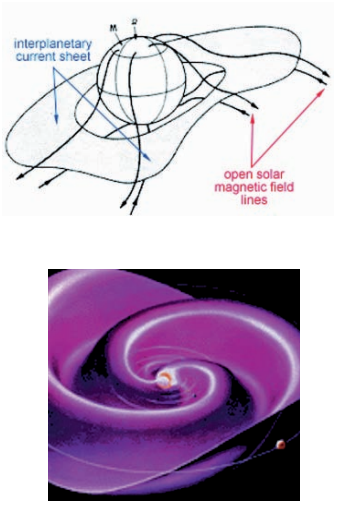


Figure 3.13: The interplanetary current sheet (above) and the heliospheric current sheet (below). Courtesy: http://pluto.space.swri.edu/image/glossary/IMF.html

The IMF is a vector. The two components B_x, B_y are oriented parallel to the ecliptic. The component B_z is perpendicular to the ecliptic. It is created by disturbances (e.g. waves,...) in the solar wind. When the IMF field lines and the geomagnetic field lines are oriented antiparallel to each other, they can reconnect. By this process energy, mass and momentum can be transferred from the IMF, solar wind, to the geomagnetic field, the strongest coupling occurs when B_z is oriented southward.

Generally, the strength of the IMF near Earth is between 1-37 nT and the average is ~ 6 nT.

3.7 Variations of the Solar Diameter

3.7.1 Relation Solar Diameter-Solar Dynamo

When measuring the solar diameter one has to take into account that the Sun is a gaseous sphere and its diameter is in principle a matter of definition. When looking at the solar limb, the decrease of the tangential optical depth from unity to essentially zero occurs over only a few hundred km which is small compared to the total solar radius. A major decrease occurs within 0.2 arcsec of both sides of the point of inflection when regarding a scan. Therefore, one can define a solar diameter this way.

Why is it important to study solar radius variations? The radiated energy of the sun comes from the nuclear energy generation (Fusion of H to He) in the deep solar interior. In the solar core at a temperature of more than 10 Million K the energy is generated by the fusion of H to He and high energetic γ ray photons are

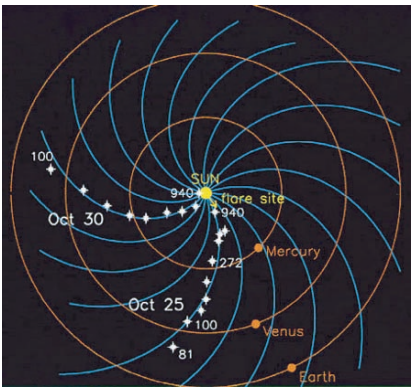


Figure 3.14: The spiral nature of the IMF. Courtesy: NASA

emitted. These energetic photons are absorbed and re-emitted in the solar interior (mean free path between the absorption processes is only 1 cm) and therefore for a photon generated by such nuclear reaction, it takes more than 1 million years to diffuse out of the core region. Thus one can argue that the emergent luminosity at the core outer boundary is effectively constant on solar cycle timescales. If there is any luminosity variability at the surface there must be an intermediate energy reservoir between the core and the photosphere. There are several mechanisms for storing energy during a solar activity cycle, such as magnetic fields or gravitational energy. Each of them leads to distinct perturbations in the equilibrium structure of the sun. Therefore, one can argue that a sensitive determination of the solar radius fluctuations can help to understand the solar cycle and it is clear that the magnitude of the radius fluctuations compared to the luminosity change contains information on where and how energy is stored.

Sofia and Endal (1979) [291] introduced the parameter W by:

$$W = \frac{\delta r}{r} / \frac{\delta L}{L} \tag{3.51}$$

The models predict a wide range for W:

- 2×10^{-4} Spruit, 1982 [298];
- 8×10^{-4} Gilliland, 1980 [110];
- 5×10^{-3} Dearborn and Blake, 1980 [72];
- 7.5×10^{-2} Sofia and Endal, 1979 [291];
- W could be positive or negative, Lydon and Sofia (1995) [204].

In the following we discuss briefly some measurement methods and give the results.

3.7.2 Ground Based Measurements

Ground based measurements of the solar diameter exist over more than 300 years. Because of the small variations the results are controversial and inconsistent. The first determination of the solar diameter was made by Aristarchus 270 BC. He obtained a value of 900 arcsec. The first accurate measurements were performed by Mouton in the year 1970 at Lyon during the period of 1959-1961 and he obtained a value of 960.6 arcsec for the solar semidiameter. From historical data it may be deduced that the solar radius may have been larger during the Maunder Minimum. As we have seen this minimum of solar activity coincided with extremely cold periods in Europe and the Atlantic regions (Ribes *et al.* 1991). Also Laclare *et al.* (1996) [182] found a larger solar radius during solar minimum. However other groups (Ulrich and Bertello, 1995 [321], Noel, 1997 [232] and Basu 1998 [29]) found a positive correlation: the solar diameter increases with enhanced solar activity. Besides a possible variation of the solar radius with the solar cycle there are also hints that the solar radius changes over timescales of 1 000 days to 80 years (Gilliland, 1980 [110]). Thus we see that there is a wide range of measurements and the results are ambiguous. From helioseismic measurements Dziembowski *et al.* (2000) [86] deduced solar radius fluctuations and they found a change of 10 mas [20] between 1996 and 1998.

Laclare *et al.* (1996) [182] published results of solar diameter measurements obtained with the Danjon astrolabe at the Observatoire de la Côte d'Azur; this program was initiated in 1975 and the instrument consists of a set of 11 reflector prisms which enables the measurement of the diameter up to 22 times a day at different zenith distances (from 30 to 70 degrees). Observing a transit requires the recording of the time when both images of the Sun's edge, i.e. the direct one and its reflection on a mercury surface become tangent to each other. At this instant the Sun's edge crosses the parallel of altitude (almucantar) which is defined in the instrument by the angle of the reflector prism and also by the refraction and other terms.

Of course this technique requires a true stability of the almucantar during observation and Zerodur types of ceramic reflector prims (which are practically unaffected by dilation) and a mercury mirror establishing the horizontal plane are used.

Furthermore, the observations were cleared of personal bias by using an acquisition system equipped with a CCD camera at the focal plane of the instrument. The limb was defined as the point where the intensity distribution on a CCD line has its inflection point (zero of the second derivative of the solar limb function). For each frame then the limb was reconstructed by a least-square adjustment of a parabola through the inflection points.

The mean value of the semi diameter was obtained by visual measurements and by the above described data acquisition system:

- 5 000 visual measurements, same observer; 1975-1994: 959.46±0.01 arcsec., broad band (200 nm) filter was used centered on 540 nm.

[20] 1 mas = 1 milli arcsec

- CCD acquisition program: 981 CCD measurements in the period 1989-1994; mean value= 959.40±0.01 arcsec.

- Correlations with solar activity: nearly opposing trend; high activity means smaller diameter.

It is important to notice that all ground based observations must take into account the quasi biennial oscillation in the Earth's atmosphere.

Other astrolabe measurements were done by Sanchez *et al.* (1995) [268] at the San Fernando Observatory (Cadiz), Noel (1995) [233] at Santiago and Leister *et al.* at Sao Paulo Observatory (1990) [193]. Ribes *et al.* (1991) [256] report on photoelectric measurements made at HAO in Boulder from 1986-1990; Wittmann *et al.* (1993 [340]) report on measurements using drift timing in Izana and Locarno. Other authors measured the solar diameter from eclipse data (e.g. Kubo, 1993) [177].

3.7.3 Satellite Measurements

The fact that these measurements are controversial is related to the problem that the fluctuations are quite small and Earth bound observations are always limited by seeing. Thus one wants to reduce this effect by using balloon borne instruments (Sofia et al. 1994) or satellite data (Michelson Doppler Imager, MDI on board of SOHO). These data are free of atmospheric disturbances and promise very accurate determinations of the solar radius.

Emilio *et al.* (2000) [89] used SOHO MDI measurements to derive possible variations of the solar diameter. They used 1 minute cadence images, and these were low pass filtered in order to remove solar 5 minute p mode intensity oscillations. The limb pixels (2 arcsec/pixel) were downlinked every 12 minutes. The data set used was between 1996 April 19 and 1998, June 24. They did not use data obtained after the recovery of SOHO in November 1998 because of the frequent instrument mode interruptions and focal length calibration difficulties. They find annual radius variations at an amplitude of 0.1 arcsec and a secular increase of about the same amplitude over the period between 1996 and 1998. The systematic variation is caused by the changing thermal environment of the MDI front window which yields small but measurable changes in the telescope focal length. A temperature gradient of a few degrees from the center of the window to the aluminium cell at the filter edge can produce a weak lens effect; that corresponds to a focal length of a few km and changes the telescope focal length by a few parts in 10^4. The secular change is also influenced by the degradation of the front window and increased absorptivity.

Thus the MDI data yield lower values of opposite sign. Since Sofia *et al.* (1979) [291] claimed that $W \sim 0.075$, solar cycle changes which affect the convective efficiency near the photosphere will have a large effect on the solar radius; the MDI measurements rule out this high value of W and suggest that solar cycle luminosity changes are not caused by superficial fluctuations in the outer layers of the Sun.

Table 3.8: Solar Diameter Measurements

Author	Period	Value	Corr. coeff.
Wittmann	1972-1991	+0.25 arcsec	0.9
Laclare	1978-1994	+0.09	0.4
Leister	1980-1993	+0.09	0.2
Kubo	1970-1991	+0.05	0.8
Bode	1976 1994	0.00	0.1
Neckel	1981-1990	0.00	0.1

The French CNES plans to operate the satellite mission PICARD which is supposed to operate in 2008 and the satellite will be launched with a Russian Dnepr rocket. One of the experiments of this low mass mission is SODISM (Solar Diameter Imager and Surface Mapper).

Chapter 4

MHD and the Solar Dynamo

In this chapter we will explain the basic MHD equations which are needed to understand solar active phenomena such as spots, prominences, flares etc. The solar dynamo is needed to maintain the solar activity cycle[1].

We sometimes mix the unit system. This was done by intention because in some cases it is easier and more intuitive to deviate from the SI system.

Let us give a simple example. The Coulomb law is defined as[2]:

$$F = k\frac{q_1 q_2}{r^2} \tag{4.1}$$

In the SI system $k = 1/4\pi\epsilon_e; \epsilon_0 = 8.854 \times 10^{-12}\,\mathrm{C^2N^{-1}m^{-2}}$. In the cgs system k=1.

4.1 Solar Magnetohydrodynamics

4.1.1 Basic Equations

To understand the surface activity of the Sun and the solar cycle it is necessary to briefly outline the principles of MHD. The properties of electromagnetic fields are described by Maxwell's equations:

$$\nabla \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial \mathbf{t}} \tag{4.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{4.3}$$

$$\mathrm{div}\mathbf{B} = 0 \tag{4.4}$$

$$\mathrm{div}\mathbf{D} = \rho_\mathrm{E} \tag{4.5}$$

Here $\mathbf{H}, \mathbf{B}, \mathbf{D}, \mathbf{E}, \mathbf{j}, \rho_\mathrm{E}$ are the magnetic field, magnetic induction, electric displacement, electric field, electric current density and electric charge density.

[1] See also e.g. Advances in Solar System Magnetohydrodynamics, E.R. Priest, A. W. Hood, Cambridge, 1991

[2] it defines the force acting between two charges q_1, q_2

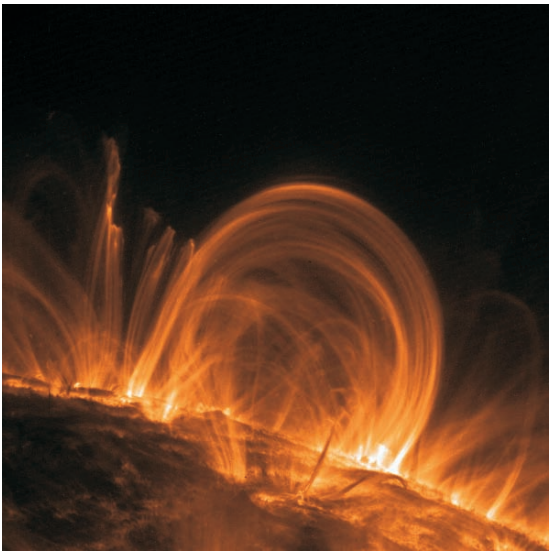


Figure 4.1: Looped magnetic field lines in the solar chromosphere and corona. Photo: TRACE mission, NASA

Equation 4.2 is the Ampere law which states that a spatially varying magnetic field (given by $\nabla \times \mathbf{H}$) produces currents- in MHD often the variation of the $\mathbf{E}$ fields is neglected, thus the term $\frac{\partial \mathbf{D}}{\partial t} \rightarrow 0$. We can also state that a current $\mathbf{j}$ induces a magnetic field that is in a direction opposite to it.

Equation 4.3 is the Faraday law and states that a time varying magnetic field produces an electric field.

If μ_0, ϵ_0 are the permeability and permittivity of free space, then for most gaseous media in the universe:

$$\mathbf{B} = \mu_\mathbf{0}\mathbf{H} \qquad \mathbf{D} = \epsilon_\mathbf{0}\mathbf{E} \tag{4.6}$$

The following equation relates the electric current density to the fields producing it (generalized Ohm's law):

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \tag{4.7}$$

σ is the electrical conductivity and $\mathbf{u}$ is the bulk velocity of the matter. The final equations depend on the state of matter; if it consists of electrons and one type of ion:

$$\mathbf{j} = n_i Z_i e \mathbf{u_i} - n_e e \mathbf{u_e} \qquad \rho_\mathrm{E} = n_i Z_i e - n_e e \tag{4.8}$$

$n_i, \mathbf{u_i}, n_e, \mathbf{u_e}$ are the number density and velocity of the ions and electrons respectively and $Z_i e, -e$ are the charges on the ion and the electron.

In astrophysics two simplifications are applied:

- magnetic fields are treated as permanent,
- electric fields are regarded as transient.

The third Maxwell equation (4.4) states that there are no magnetic monopoles[3]. Electric fields can be produced by separating positive and negative charges through the fourth Maxwell equation (4.5) however the attraction between these charges is so strong that charge separation is usually cancelled out very quickly. Through the second Maxwell equation electric fields can be produced by time varying magnetic fields. Such fields are only significant, if there are rapid changes by time varying magnetic fields. Magnetic fields produced by the displacement current $\partial\mathbf{D}/\partial t$ are usually insignificant in astrophysical problems because electric fields are unimportant; however they can be produced by a conduction current $\mathbf{j}$, if the electrical conductivity is high enough. Such magnetic fields may be slowly variable in time and space.

We therefore neglect $\partial\mathbf{D}/\partial\mathbf{t}$, combine the equations

$$\nabla \times \mathbf{H} = \mathbf{j} \qquad \nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial\mathbf{t} \qquad \mathbf{B} = \mu_\mathbf{0}\mathbf{H} \qquad \mathbf{j} = \sigma\mathbf{E}, \tag{4.9}$$

obtaining

$$\frac{\partial \mathbf{B}}{\partial t} + \frac{1}{\mu_o \sigma} \nabla \times \nabla \times \mathbf{B} = \mathbf{0} \tag{4.10}$$

and using $\nabla \times \nabla \times \mathbf{B} = \mathrm{grad}\,\mathrm{div}\mathbf{B} - \nabla^\mathbf{2}\mathbf{B}$ and $\mathrm{div}\mathbf{B} = \mathbf{0}$:

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B} \tag{4.11}$$

This is also called the *induction equation*- for the static case. In cartesian coordinates this equation for the x coordinate is:

$$\frac{\partial B_x}{\partial t} = \frac{1}{\mu_0 \sigma} \left[\frac{\partial^2 B_x}{\partial x^2} + \frac{\partial^2 B_y}{\partial y^2} + \frac{\partial^2 B_z}{\partial z^2} \right] \tag{4.12}$$

The solution of these equations shows that magnetic fields decay together with the current producing them. We can derive an approximate decay time: let us assume that currents vary significantly in distance L, then from (4.11) the decay time becomes

$$\tau_\mathrm{D} = \mu_0 \sigma L^2 \tag{4.13}$$

If at time $t = 0$ there exists a sinusoidal field

$$B_x = B_0 \exp(iky) \tag{4.14}$$

the solution at a later time t is:

$$B_x = B_0 \exp(iky) \exp(-k^2 t/\mu_0 \sigma) \tag{4.15}$$

[3]This is a common experience: a division of a permanent magnet into two does not separate north and south poles.

The wavelength λ of the spatial variation of the field is $2\pi k$, the original field decays by a factor e in the time $\mu_0\sigma\lambda^2/4\pi^2$.

Let us consider typical fields of stars: the dimension of the star L and the electrical conductivity are both high (if the gas is fully ionized). Therefore, the lifetime of a magnetic field could exceed the main sequence lifetime, such a field is called a *fossil field*.

The same is not true for the Earth. Its field is produced by currents in a liquid conducting core and continuously regenerated by a *dynamo mechanism*.

The electrical conductivity of an ionized gas is $\sim T^{3/2}$. That means that the characteristic time for decay of currents in the outer layers of the Sun is much less than the solar lifetime, whereas the decay near the center exceeds the lifetime (since the temperature near the surface is about 6 000 K and near the center about 1.5×10^7 K). If the field in the solar interior were a fossil field extending throughout the Sun, the field in the outer layers would now be current free - similar to the field of a dipole. However we don't observe this. The surface field is very complex and therefore it must be also regenerated by a dynamo. It is conceivable that a fossil field of the Sun was destroyed at the very early evolution of the Sun, when it was fully convective before reaching the main sequence. Also helioseismology argues against a strong field.

4.1.2 Some Important MHD Effects

A magnetic field in a conducting fluid exerts a force per unit volume which is

$$\mathbf{F}_{\rm mag} = \mathbf{j} \times \mathbf{B} = ({\rm curl}\mathbf{B} \times \mathbf{B})/\mu_0 = -{\rm grad}(B^2/2\mu_0) + \mathbf{B}.\nabla\mathbf{B}/\mu_0 \qquad (4.16)$$

This can be interpreted as:

- ${\rm grad}(B^2/2\mu_0)$ isotropic pressure,
- $\mathbf{B}.\nabla\mathbf{B}/\mu_\mathbf{0}$ tension along the lines of magnetic induction.

The isotropic pressure must be added to the gas pressure: let us assume we have a tube of magnetic flux, and $P_{\rm out}$ denotes the pressure outside and $P_{\rm in}$ the pressure inside the tube, then for equilibrium:

$$P_{\rm out} = P_{\rm in} + B^2/2\mu_0 \qquad (4.17)$$

The gas pressure can be written as $\Re\rho T/\mu$, where $\Re$ is the gas constant and μ the mean molecular weight. With $T_{\rm in} = T_{\rm out}$ we must have:

$$\rho_{\rm in} < \rho_{\rm out} \qquad (4.18)$$

A tube of magnetic flux is lighter than its surroundings and will start to rise which is called *magnetic buoyancy*.

To a good approximation, the fluid is tied to the magnetic field. For the Sun two extremes occur:

- photosphere: the fluid motions drag the magnetic field lines around (the magnetic field is frozen in);

- corona: the magnetic force is so strong that it constrains the motion of the fluid.

The tying of the fluid to the magnetic field lines also permits the propagation of MHD waves which have some similarity to sound waves but a characteristic speed (Alfvén speed):

$$c_{\mathrm{H}} = \sqrt{B^2/\mu_0\rho} \tag{4.19}$$

The sound speed is given by

$$c_{\mathrm{s}} = \sqrt{\gamma P/\rho} \tag{4.20}$$

This can also be seen from the induction equation. Let us consider again the Maxwell equations. From $\mathbf{j} = \sigma\mathbf{E} + \mathbf{u} \times \mathbf{B}$ we can extract $\mathbf{E}$:

$$\mathbf{E} = \frac{\mathbf{j}}{\sigma} - \mathbf{u} \times \mathbf{B} \tag{4.21}$$

This is substituted into the Maxwell equation (4.3) yielding:

$$\nabla \times \left(\frac{\mathbf{j}}{\sigma} - \mathbf{u} \times \mathbf{B} \right) = -\frac{\partial \mathbf{B}}{\partial t} \tag{4.22}$$

We have already argued that the displacement current can be neglected in the first Maxwell equation and therefore $\nabla \times \mathbf{B} = \mu\mathbf{j}$, from which $\mathbf{j} = \mathbf{1}/\mu\nabla \times \mathbf{B}$ and

$$\nabla \times \left(\frac{1}{\mu\sigma}\nabla \times \mathbf{B} - \mathbf{u} \times \mathbf{B} \right) = -\frac{\partial \mathbf{B}}{\partial t} \tag{4.23}$$

using the formula

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \mathbf{A}) - \boldsymbol{\nabla^2 \mathbf{A}} \tag{4.24}$$

from vectoranalysis, we get:

$$\nabla \times \left(\frac{1}{\mu\sigma}\nabla \times \mathbf{B} \right) = \frac{1}{\mu\sigma} \left[\nabla(\nabla \mathbf{B}) - \boldsymbol{\nabla^2 \mathbf{B}} \right]$$

This gives us the final form of the so called *induction equation*:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \eta \boldsymbol{\nabla^2 \mathbf{B}} \tag{4.25}$$

Here $\eta = 1/\mu\sigma$ is the magnetic diffusivity. The case where the plasma is stationary was already discussed above ($\mathbf{u} = \mathbf{0}$), the field decays in the ohmic decay time $\tau = L^2/\eta$. Let us discuss the case when $\eta = 0$. Then, the field $\mathbf{B}$ is completely determined by the plasma motions $\mathbf{u}$ and the induction equation is the equivalent to the vorticity equation for an inviscid fluid. The magnetic flux Φ through a material surface S which is a surface that moves with the field, is:

$$\Phi = \int_S \mathbf{B}.\mathbf{dS} \tag{4.26}$$

If G is the material closed curve bounding S, the total rate of change of Φ is (see also eq. 4.34):

$$\frac{D\Phi}{Dt} = \int_S \frac{\partial \mathbf{B}}{\partial t}.d\mathbf{S} + \oint_{\mathbf{G}} \mathbf{B}.(\mathbf{u} \times \mathbf{dl}) \tag{4.27}$$

$$= \int_S \frac{\partial B}{\partial t}.d\mathbf{S} + \oint_{\mathbf{G}} (\mathbf{B} \times \mathbf{u}).\mathbf{dl} \tag{4.28}$$

$$= \int_S \left[\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{u} \times \mathbf{B}) \right] .d\mathbf{S} \tag{4.29}$$

In the last equation we have used the Stokes Theorem:

$$\int_L (A)\mathbf{Cdl} = \int_{\mathbf{A}} (\nabla \times \mathbf{C})\mathbf{dA}' \tag{4.30}$$

If $\eta = 0$ the total flux across any arbitrary surface moving with the fluid remains constant, the magnetic field lines are said to be *frozen* in to the flow.

If v_0, l_0 are typical velocity and length-scale values for our system, then the ratio of the two terms on the right hand side of the induction equation gives the *Magnetic Reynolds Number*

$$R_m = l_0 v_0 / \eta_0 \tag{4.31}$$

In an active solar surface region one has $\eta_0 = 1\,\mathrm{m}^{-2}\mathrm{s}^{-1}$, $l_0 = 700\,\mathrm{km} \sim 1\,\mathrm{arcsec}$ and $v_0 = 10^4\,\mathrm{m/s}$ we find $R_m = 7 \times 10^9 >> 1$. Thus the field is frozen to the plasma and the electric field does not drive the plasma but is simply $\mathbf{E} = -\mathbf{u} \times \mathbf{B}$. However, if the length-scales of the system are reduced the diffusion term $\eta \nabla^2 \mathbf{B}$ becomes important. Then the field lines are allowed to diffuse through the plasma and this yields to magnetic braking and changing the global topology of the field (magnetic reconnection).

4.1.3 Magnetic Reconnection

Magnetic reconnection is the process by which lines of magnetic force break and rejoin in a lower energy state. The excess energy appears as kinetic energy of the plasma at the point of reconnection. In Fig. 4.2 single arrow lines denote magnetic field and double line arrows the magnetofluid velocity. As it can be seen, the merging of two magnetofluids with oppositely oriented magnetic fields causes the field to annihilate. The excess energy accelerates the plasma out of the reconnection region in the direction of the full double line arrows. Note the characteristic X-point, where the topology changes for the field lines.

The plasma, where the field is annihilated is accelerated outwards to Alfvén speed v_A:

$$v_A = B_0 / \sqrt{4\pi M n_B} \tag{4.32}$$

n_B... density inside the current sheet, M the plasma average molecular weight.

A similar process occurs in coronal loops that were observed in hard and soft x-rays by Yohkoh and SOHO instruments. Such a coronal loop (see right drawing in Fig. 4.2) is stretched out by pressure which is provided by buoyancy. A magnetic

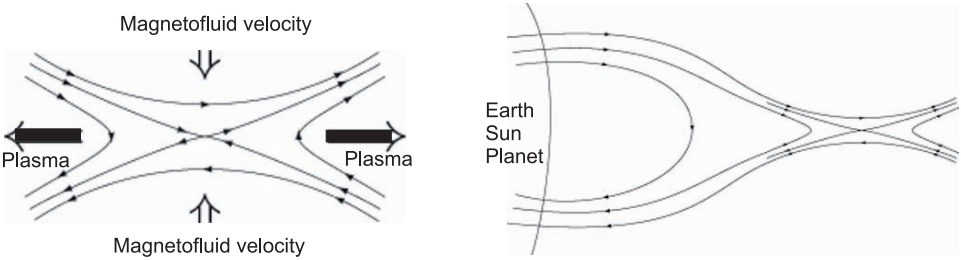


Figure 4.2: Principle of magnetic reconnection.

structure is buoyant because the particle density is lower there since it contains larger magnetic energy density (see magnetic buoyancy). Thus the external pressure is balanced by a lower gas pressure in conjunction with a magnetic pressure. The top of the loop distends and reconnection occurs. Particles in the reconnection region accelerate towards the surface of the sun and out away. Those particles that are accelerated towards the sun are confined within the loop's magnetic field lines and follow these lines to the footpoint of the loop where they collide with other particles and lose their energy through x-ray emissions. Such processes are the cause of solar flares and will be discussed in the next chapter.

Magnetic reconnection also provides a mechanism for energy to be transported into the solar corona.

A similar process occurs in the earth's magnetotail. The solar wind distends the Earth's dipole field so that the field extends far behind the Earth. Earthward flowing plasma streams with flow velocities up to 1000 km/s (which is close to the local Alfvén speed) have been observed (Birn *et al.* 1981 [37]).

A recent review on solar MHD was given by Walsh (2001) [329].

4.1.4 Fluid Equations

The continuity or mass equation for a fluid is:

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\nabla.\mathbf{u} = \mathbf{0} \tag{4.33}$$

and the total derivative means here:

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \mathbf{u}.\nabla \tag{4.34}$$

(See any textbook on fluid dynamics for a derivation of this formula). Now let us consider the equation of motion in a plasma with velocity $\mathbf{u}$: the momentum equation includes the Lorentz force term $\mathbf{j} \times \mathbf{B}$ and other forces $\mathbf{F}$, such as gravity and viscous forces:

$$\rho\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\nabla p + \mathbf{j} \times \mathbf{B} + \mathbf{F} \tag{4.35}$$

Here p is the plasma pressure. Let us assume a Newtonian fluid with isotropic viscosity, then $\mathbf{F}$ may be written as:

$$\mathbf{F} = -\rho \mathbf{g}(\mathbf{r}) \frac{\mathbf{r}}{\mathbf{r}} + \rho \nu \nabla^{\mathbf{2}} \mathbf{u} \tag{4.36}$$

$g(r)$ is the local gravity acting in the radial direction and ν the kinematic viscosity. Let us make thinks more complicated: Consider a frame of reference with angular velocity Ω at a displacement $\mathbf{r}$ from the rotation axis:

$$\rho \frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t} = -\nabla p + \mathbf{j} \times \mathbf{B} + \mathbf{F} + \rho \left[2\mathbf{u} \times \mathbf{\Omega} + \mathbf{r} \times \frac{d\mathbf{\Omega}}{dt} + \frac{1}{2} \nabla |\mathbf{\Omega} \times \mathbf{r}|^2 \right] \tag{4.37}$$

The three terms in [ ] denote: Coriolis force, change of rotation and centrifugal force. Stars rotate more rapidly when they are young. Under most circumstances the latter two terms are small compared with the Coriolis term $\mathbf{u} \times \mathbf{\Omega}$.

4.1.5 Equation of State

The perfect gas law

$$p = \frac{k \rho T}{m} = nkT \tag{4.38}$$

determines the constitution of stars, $k = 1.38 \times 10^{-23}$ J/K being the Boltzmann constant, m is the mean particle mass and n the number of particles per unit volume. If s denotes the entropy per unit mass of the plasma, then the flux of energy (heat) through a star becomes:

$$\rho T \frac{\mathrm{D}s}{\mathrm{D}t} = -L \tag{4.39}$$

L is the energy loss function. This function describes the net effect of all the sinks and sources of energy. For MHD applications this becomes:

$$\frac{\rho^\gamma}{\gamma - 1} \frac{\mathrm{D}}{\mathrm{D}t} \left(\frac{p}{\rho^\gamma} \right) = -\nabla . \mathbf{q} + \kappa_r \nabla^2 T + \frac{j^2}{\sigma} + H \tag{4.40}$$

In this equation we have:

- $\mathbf{q}$: heat flux due to conduction
- κ_r: coefficient of radiative conductivity
- T temperature
- j^2/σ ohmic dissipation (Joule heating)
- H represents all other sources.

4.1.6 Structured Magnetic Fields

If the plasma velocity is small compared with the sound speed ($\sqrt{\gamma p/\rho}$), the Alfvén speed ($\sqrt{B/\mu\rho}$) and the gravitational free fall speed ($\sqrt{2gl}$), the inertial and viscous terms in equation 4.35 may be neglected yielding:

$$0 = -\nabla p + \mathbf{j} \times \mathbf{B} + \mathbf{F} \tag{4.41}$$

This equation must then be solved with $\nabla \times \mathbf{B} = ...$, $\nabla.\mathbf{B} = \mathbf{0}$ and the ideal gas law as well as a simplified form of the energy equation.

Let us introduce the concept of *scale height*. Let

$$0 = -\frac{dp}{dz} - \rho g \tag{4.42}$$

Substitute in the above equation $\rho = pm/kT$ (ideal gas) and integrate:

$$p = p_0 \exp\left[-\int_0^z \frac{dz}{H_p(z)}\right] \tag{4.43}$$

(p_0 is the pressure at $z = 0$). This defines the local pressure scale height H_p:

$$H_P = kT/mg = p/\rho g \tag{4.44}$$

At solar photospheric temperatures ($T \sim 5000\,\mathrm{K}$) we find $H_p = 0.150\,\mathrm{Mm}$, whereas at coronal temperatures $T \sim 10^6\,\mathrm{K}$ we find $H_p \sim 30\,\mathrm{Mm}$.

That concept can also be applied to MHD in the case of magnetostatic balance discussed above. Assume that gravity acts along the negative z direction and s measures the distance along the field lines inclined at angle θ to this direction, then the component of eq. (4.41) in the z-direction becomes:

$$0 = -\frac{dp}{ds} - \rho g \cos\theta \qquad dz = ds\cos\theta \tag{4.45}$$

Therefore, the pressure along a given field line decreases with height, the rate of decrease depends on the temperature structure (given by the energy equation).

If the height of a structure is much less than the pressure scale height, gravity may be neglected. The ratio β is given by gas pressure p_0 to magnetic pressure $B_0^2/2\mu$. If $\beta << 1$, any pressure gradient is dominated by the Lorentz force and (4.41) reduces to:

$$\mathbf{j} \times \mathbf{B} = \mathbf{0} \tag{4.46}$$

In this case the magnetic field is said to be *force free*. In order to satisfy (4.46) either the current must be parallel to $\mathbf{B}$ (Beltrami fields) or $\mathbf{j} = \nabla \times \mathbf{B} = \mathbf{0}$. In the latter case the field is a current free or potential field.

If β is not negligible and the field is strictly vertical of the form $\mathbf{B} = \mathbf{B}(\mathbf{x})\mathbf{k}$, then (4.41) becomes:

$$0 = \frac{\partial}{\partial x}\left[p + \frac{B^2}{2\mu}\right] \tag{4.47}$$

4.1.7 Potential Fields

Potential fields result when $\mathbf{B}$ vanishes. We can write $\mathbf{B} = \nabla\mathbf{A}$ so that $\nabla \times \mathbf{B} = \mathbf{0}$; with $\nabla.\mathbf{B} = \mathbf{0}$ one obtains Laplace's equation:

$$\nabla^2\mathbf{A} = \mathbf{0} \tag{4.48}$$

If the normal field component B_n is imposed on the boundary S of a volume V, then the solution within V is unique. Also if B_n is imposed on the boundary S, then the potential field is the one with the minimum magnetic energy.

These two statements have many implications for the dynamics of the solar atmosphere. During a solar flare e.g. the normal field component through the photosphere remains unchanged. However, since enormous amounts of energy are released during the eruptive phase, the magnetic configuration cannot be potential. The excess magnetic energy could arise from a sheared force-free field.

Let us consider an example of a potential field in two dimensions: Consider the solutions $A(x, z) = X(x)Z(z)$ such that $\nabla^2 A = 0$ gives:

$$\frac{1}{X}\frac{d^2X}{dx^2} = -\frac{1}{Z}\frac{d^2Z}{z^2} = -n^2 \tag{4.49}$$

where $n = const.$ A solution to (4.49) would be:

$$A = \left(\frac{B_0}{n}\right)\sin(nx)e^{-nz} \tag{4.50}$$

this gives for the field components:

$$B_x = \frac{\partial A}{\partial x} = B_0\cos(nx)e^{-nz} \tag{4.51}$$

$$B_z = \frac{\partial A}{\partial z} = -B_0\sin(nx)e^{-nz} \tag{4.52}$$

The result is a two dimensional model of a potential arcade.

4.1.8 3 D Reconstruction of Active Regions

If we look at an active region on the solar disk center we have no information about the 3 D structure of it, especially about the 3 D magnetic field configuration which is important for modelling such regions. Information about the height dependence of active regions can only be obtained when observing such features near the solar limb. Let us consider some simple model to reconstruct these features.

$$B_x = \frac{\partial A}{\partial z},\ B_y(x, z),\ B_z = -\frac{\partial A}{\partial x} \tag{4.53}$$

We see immediately that $\nabla.\mathbf{B} = \mathbf{0}$ Let us assume that the footpoints of the field are anchored down into the photosphere (z=0). Projecting the resulting field onto the xz plane gives:

$$\frac{\partial A}{\partial x}dx + \frac{\partial A}{\partial z}dz = 0 \tag{4.54}$$

Therefore $dA = 0, A = const.$ From $\nabla \times \mathbf{B} = \mu\mathbf{j}$ calculate the components of the current density:

$$j_x = -\frac{1}{\mu}\frac{\partial B_y}{\partial z} \tag{4.55}$$

$$j_y = \frac{1}{\mu}\left(\frac{\partial B_y}{\partial z} - \frac{\partial B_y}{\partial x}\right) \tag{4.56}$$

$$j_z = \frac{1}{\mu}\frac{\partial B_y}{\partial x} \tag{4.57}$$

And then the components of the Lorentz force ($\mathbf{j} \times \mathbf{B}$):

$$\nabla^2 A\frac{\partial A}{\partial x} + B_y\frac{\partial B_y}{\partial x} = 0 \tag{4.58}$$

$$\frac{\partial B_y}{\partial x}\frac{\partial A}{\partial z} - \frac{\partial A}{\partial x}\frac{\partial B_y}{\partial z} = 0 \tag{4.59}$$

$$\nabla^2 A\frac{\partial A}{\partial z} + B_y\frac{\partial B_y}{\partial z} = 0 \tag{4.60}$$

4.1.9 Charged Particles in Magnetic Fields

In this chapter we consider first the motion of a single charged particle in a given electromagnetic field. The particle has charge q and the equation of motion is:

$$m\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \tag{4.61}$$

Let us write:

$$\begin{aligned} \mathbf{B_0} &= B_0\mathbf{b} \\ \mathbf{E_0} &= E_\parallel\mathbf{b} + \mathbf{E}_\perp \\ \mathbf{v} &= v_\parallel\mathbf{b} + \mathbf{v}_\perp \end{aligned}$$

where $\mathbf{b}$ is a unit vector and $\mathbf{v} \times \mathbf{B_0} = \mathbf{B_0}(\mathbf{v}_\perp \times \mathbf{b})$ is perpendicular to $\mathbf{b}$. Equation 4.61 splits into a parallel and a perpendicular component (e.g. parallel means parallel to the magnetic field lines):

$$m\frac{\mathrm{d}v_\parallel}{\mathrm{d}t} = qE_\parallel \tag{4.62}$$

$$m\frac{dv_\perp}{dt} = q[E_\perp + B_0(v_\perp \times \mathbf{b})] \tag{4.63}$$

Equation 4.62 has the solution

$$v_\parallel = (qE_\parallel/m)t + v_{\parallel 0} \tag{4.64}$$

Here $v_{\parallel 0}$ is the velocity component at $t = 0$ in the direction of the magnetic field line. We see that particles of opposite sign of charge q move in opposite directions, they move along an electric field parallel to a magnetic field which destroys $E_\parallel$.

Now let us solve equation 4.63 by writing:

$$\mathbf{v}_\perp = \mathbf{v}'_\perp + \mathbf{E}_\perp \times \mathbf{b}/\mathbf{B_0} \tag{4.65}$$

and equation 4.63 becomes:

$$m\frac{\mathrm{d}\mathbf{v}'_\perp}{\mathrm{d}t} = qB_0(\mathbf{v}'_\perp \times \mathbf{b}) \tag{4.66}$$

Summarizing we arrive at:

- Motion in which acceleration is $\perp$ to the velocity,
- const. acceleration to the velocity.

This is a motion in a circle around the direction of $\mathbf{b}$ and the motion has frequency $|q|B_0/m$, the magnitude of the velocity is $v_{\perp 0}$; the radius of the orbit, the gyration radius r_g is

$$r_g = mv_{\perp 0}/|q|B \tag{4.67}$$

For an electron the gyration frequency is 1.8×10^{11}(B/Tesla)Hz. The corresponding gyration radius is $6 \times 10^{-9}(v_{\perp 0}/\,\mathrm{km/s})(B/\mathrm{Tesla})\mathrm{m}$.

The difference in mass between electrons and protons is about 1800. What follows for the radius and frequency of gyration?

The gyration follows also from the simple statement, that in the absence of other forces, the Lorentz force balances the centripetal force of the particle's motion around the field line. Let α be the pitch angle, which is the angle between the direction of motion and the local field line. Then for the gyro radius (Larmor radius :

$$R_L = \frac{cp_\perp}{qB} = \left|\frac{mc\mathbf{v} \times \mathbf{B}}{qB^2}\right| \tag{4.68}$$

The Larmor radius for a 100 keV electron (which is typical for electrons in the inner radiation belt of the Earth) is about 100 m.

Let us assume, that the magnetic flux through a particle's orbit is constant- this is certainly the case when changes of the magnetic field are small over the gyro radius and one gyro period. From the condition that $d\Phi_B/dt = 0$ the so called first adiabatic invariant follows:

$$\mu_B = \frac{p_\perp^2}{2mB} \tag{4.69}$$

or in terms of the particle's energy:

$$\mu_B = \frac{E\sin^2\alpha}{B} \tag{4.70}$$

From the conservation of the first adiabatic invariant it follows, that the pitch angle increases, when the particle moves to larger field strength, until $\alpha = 90^0$ at the mirror point.

Summarizing the motion of a particle:

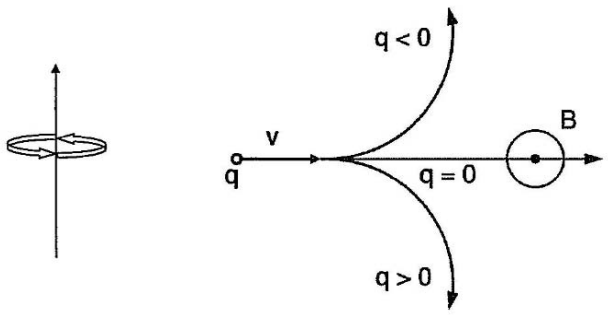


Figure 4.3: Left: Gyration of a charged particle around magnetic field lines, the sense depends on the + or - charge; right: charged particles are deflected by magnetic fields, depending on their charge. Here the magnetic fieldlines point vertically inwards, denoted by $\odot$

- accelerated motion along the field lines,
- circular motion around the field,
- drift velocity $\mathbf{E}_\perp \times \mathbf{b}/\mathbf{B_0}$ perpendicular to both electric and magnetic fields,
- the sense of the accelerated and the circular motions depends on the sign of the electric charge.
- the drift velocity is the same for all particles,
- in the absence of electric fields, a particle moves with a constant velocity in the direction of the magnetic field and with a velocity of constant magnitude around the field, thus it moves along a helical path.
- In all this discussion we have neglected one important effect. Accelerated charged particles radiate, for non relativistically moving particles this radiation is known as *cyclotron radiation* and for relativistic particles as *synchrotron radiation*.

Finally, if there is a constant non magnetic force $\mathbf{F}$ perpendicular to $\mathbf{B}$, there is a drift velocity:

$$\mathbf{v}_{\mathrm{DF}} = \frac{c\mathbf{F} \times \mathbf{B}}{qB^2} \tag{4.71}$$

Please note that again $\mathbf{v}_{\mathrm{DF}}$ is charge dependent.

Let us give some examples for drift motions:

- Gradient $\mathbf{B}$ drift: the field strength in planetary magnetospheres decreases with increasing distance from the planet and the gradient in the field strength induces a force which can be written as:

$$\mathbf{F} = -\mu_B \nabla B \tag{4.72}$$

and the drift velocity:

$$\mathbf{v_B} = \frac{\mu_B c \mathbf{B} \times \nabla B}{qB^2} \tag{4.73}$$

→ Particles move perpendicular to the lines of force and perpendicular to the magnetic gradient. Protons and electrons drift in opposite direction. In the Earth's magnetic field, electrons drift in the eastward direction. This drift motion causes a current system known as ring current. The ring current strengthens the field on its outside, helping expand the size of the magnetosphere but weakens the magnetic field in its interior. The ring current plasma population is enhanced in a magnetic storm. A southward oriented IMF (Bz negative) leads to reconnection at the side of the magnetosphere to the sun and the magnetopause is pushed closer to the Earth. Also injection of particles from the tail occurs. This enhances the number of particles. For a typical 100 keV electron or proton it takes about 5-6 hours to complete one orbit drift.

- field line curvature drift: particles move along curved field lines. The guiding center follows the curved field line and the resulting centripetal force is equal to:

$$\mathbf{F_c} = \frac{m\mathbf{v_\parallel^2}}{R_c}\mathbf{n} \tag{4.74}$$

where $\mathbf{n}$ is a unit vector outwards. The drift motion is perpendicular to the field line's radius of curvature and the field line itself.

- gravitational field drift. Let us assume that $\mathbf{F}$ is the gravitational force $\mathbf{F} = m\mathbf{g}$, then

$$\mathbf{v}_{\mathrm{DF}} = m\mathbf{g} \times \mathbf{B}/qB^2 \tag{4.75}$$

Thus the drift velocity depends on the mass/charge ratio, the ion drift is much larger than the electron drift; the particles drift in opposite directions, a current is produced.

- electric field drift: here the force is

$$\mathbf{F} = q\mathbf{E} \tag{4.76}$$

And the drift velocity

$$\mathbf{v}_E = \frac{c\mathbf{E} \times \mathbf{B}}{B^2} \tag{4.77}$$

Thus, charged particles move in a direction perpendicular to a) $\mathbf{E}$ and b) $\mathbf{B}$. Protons and electrons move in the same direction.

We stress here, that the $\nabla\mathbf{B}$ and $\mathbf{E} \times \mathbf{B}$ drift and curvature forces dominate the drift motions of particles in a magnetosphere.

Let us consider a large assembly of particles; these particles interact which is called collision. If τ_{c} is the characteristic time between collisions the collision frequency is $\nu_{\mathrm{c}} = 1/\tau_{\mathrm{c}}$. If ν_{c} is large, the particle motions will be disordered and decoupled from the magnetic field, the fluid will not be tied to the field. If

collisions are relatively rare, not only individual particles but the whole fluid will be tied to the field. The collisions provide the electrical resistivity of matter; in a fully ionized gas a good approximation to the value of the electrical conductivity is:

$$\sigma = n_e e^2 \tau_c / m_e \tag{4.78}$$

and $\tau_c \sim T^{3/2}$.

4.1.10 MHD Waves

The equation that describes the connections between the force exerted by the magnetic field and the fluid motions is

$$\rho \frac{d\mathbf{v}}{dt} = -\text{grad}P + \mathbf{j} \times \mathbf{B} + \rho \text{grad}\phi \tag{4.79}$$

The forces on the gas are the gas pressure P, the gravitational potential ϕ and the magnetic force $\mathbf{j} \times \mathbf{B}$. For a full description of the system we write down two additional equations:

a) equation of continuity (conservation of mass)[4]

$$\frac{d\rho}{dt} + \rho \text{div}\mathbf{v} = \mathbf{0} \tag{4.80}$$

b) The relation between P and ρ e.g. in the adiabatic form

$$\frac{1}{P}\frac{dP}{dt} = \frac{\gamma}{\rho}\frac{d\rho}{dt} \tag{4.81}$$

Consider the simplest case: a medium with uniform density ρ_0, pressure P_0, containing a uniform magnetic field $\mathbf{B_0}$. We ignore the influence of the gravitational field and assume that σ is so large that $\mathbf{E} + \mathbf{v} \times \mathbf{B} = \mathbf{0}$. Now let us assume a perturbation for any variable in the form of:

$$f_1 \sim \exp i(\mathbf{k}.\mathbf{r} - \omega t) = \exp i(k_x x + k_y y + k_z z - \omega t) \tag{4.82}$$

$\mathbf{k}$ is the wave vector, ω the wave frequency. The dispersion relation between ω and $\mathbf{k}$ when in the absence of a magnetic field is:

$$\omega^2 = k^2 c_s^2 \tag{4.83}$$

Therefore, in that case only one type of waves can propagate – sound waves. The wave propagates through the fluid at the wave speed $c_s = \omega/k$, $k = |\mathbf{k}|$, which is called the phase velocity of the wave.

If there is a magnetic field, the force $\mathbf{j} \times \mathbf{B}$ couples to the equation and also the Maxwell equations must be taken into account. It is very important to note that the magnetic field introduces a preferred direction into the system. In a uniform medium, sound waves travel equally strongly in all directions from its source, this is not true for MHD waves. If we write the magnetic field again in the form $\mathbf{B_0} = \mathbf{B_0 b}$, then we find three types of MHD waves:

[4]Note that d/dt is the rate of change with time following a fluid element moving with velocity $\mathbf{v}$: $\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{v}.\text{grad}$

- Alfvén waves : the dispersion relation is given by

$$\omega^2 = (\mathbf{k}.\mathbf{b})^2 c_{\rm H}^2 \tag{4.84}$$

- fast and slow magnetosonic waves; their dispersion relation is given by:

$$\omega^4 - \omega^2 k^2 (c_{\rm s}^2 + c_{\rm H}^2) + k^2 (\mathbf{k}.\mathbf{b})^2 c_{\rm s}^2 c_{\rm H}^2 = 0 \tag{4.85}$$

Let us consider two special cases: if the waves propagate along the field $\mathbf{k}.\mathbf{b} = \mathbf{k}$, there are two waves with $\omega^2 = k^2 c_{\rm H}^2$ and the sound wave $\omega^2 = k^2 c_{\rm s}^2$ unaffected by the field. For wave propagation perpendicular to the field only one wave survives with $\omega^2 = k^2(c_{\rm s}^2 + c_{\rm H}^2)$. When waves propagate anisotropically, it is necessary to introduce another wave velocity in addition to the phase velocity, the group velocity, given by $\partial\omega/\partial\mathbf{k}$ with which the wave carries energy or information. The group velocity for Alfvén waves is always $c_{\rm H}\mathbf{b}$. What does that mean? Regardless of the direction in which it propagates, energy always travels along the field lines with speed $c_{\rm H}$.

4.1.11 Magnetic Fields and Convection

Let L be the length of a box, v a typical velocity. The magnetic diffusivity is $1/\mu_0\sigma$, the eddy turnover time L/v. The resistive decay time is then $\mu_0\sigma L^2$ and the resistive decay time/eddy turnover time is denoted as magnetic Reynolds number for the flow.

$$R_{\rm m} = Lv\mu_0\sigma \tag{4.86}$$

If the magnetic Reynolds number $R_{\rm m}$ is very low, the field is unaffected by the motions, if it is high, it is wound up many times before dissipation occurs. For an intermediate value of $R_{\rm m}$ the magnetic field is carried from the center of the eddy becoming concentrated in flux ropes at the edge. This buoyant flux ropes rise towards the surface and this leads to the appearance of sunspots. However we must also take into account that convection involves different length scales. Large eddies affect the overall structure of the magnetic field as it has been just described. Others may be influenced e.g. granulation. Granulation is suppressed in a sunspot. As it was shown earlier, in the absence of magnetic field convection occurs in a gas, if the ratio of the temperature gradient to the pressure gradient satisfies the relation:

$$\frac{P}{T}\frac{{\rm d}T}{{\rm d}P} > \frac{\gamma - 1}{\gamma} \tag{4.87}$$

If a vertical magnetic field of strength B threads the fluid, then this has to be modified to:

$$\frac{P}{T}\frac{{\rm d}T}{{\rm d}P} > \frac{\gamma - 1}{\gamma} + \frac{B^2}{B^2 + \gamma P} \tag{4.88}$$

Thus a strong magnetic field can prevent convection and a weaker field can interfere with convection. Note also that the magnetic field cannot prevent motions which are oscillatory up and down the field lines but these are likely to be less efficient at carrying energy.

4.2 The Solar Dynamo

So far we have discussed the different aspects of solar activity. In the section on MHD it was shown that due to dissipation, such recurrent phenomena on the solar surface and atmosphere cannot be explained by just assuming a fossil magnetic field of the Sun. Therefore, many attempts had been made in order to explain the recurrent solar activity phenomena such as sunspots, their migration toward the equator in the course of an activity cycle etc. In the first section of this paragraph we will give a general description of the basic dynamo mechanism, in the following chapter some formulas are given.

4.2.1 The Solar Dynamo and Observational Features

Let us briefly recall what are the observational facts that a successful model for the solar dynamo must explain:

- 11 year period of the sunspot cycle; not only the number of sunspots varies over that period but also other phenomena such as the occurrence of flares, prominences,.... etc.
- the equator-ward drift of active latitudes which is known as Spörers law and can be best seen in the butterfly diagram. At the beginning of a cycle active regions appear at high latitudes and toward the end they occur near the equator.
- Hale's law: as we have mentioned the leader and the follower spot have opposite polarities. This reverses after 11 years for each hemisphere so that the magnetic cycle is in fact 22 years.
- Sunspot groups have a tilt towards the equator (this is sometimes also called Joy's law).
- Reversal of the polar magnetic fields near the time of the cycle maximum.

As we know from fundamental physics, magnetic fields are produced by electric currents. How are these currents generated in the Sun? The solar plasma is ionized and it is not at rest. There are flows on the solar surface as well as in the solar interior producing magnetic fields which contribute to the solar dynamo.

4.2.2 The $\alpha - \omega$ Dynamo

The ω Effect

Let us consider magnetic fields inside the Sun. There the conditions require that the field lines are driven by the motion of the plasma. Therefore, magnetic fields within the Sun are stretched out and wound around the Sun by differential rotation (the Sun rotates faster at the equator than near the poles). Let us consider a north-south orientated magnetic field line. Such a field line will be wrapped once around the Sun in about 8 months because of the Sun's differential rotation (Fig. 4.4).

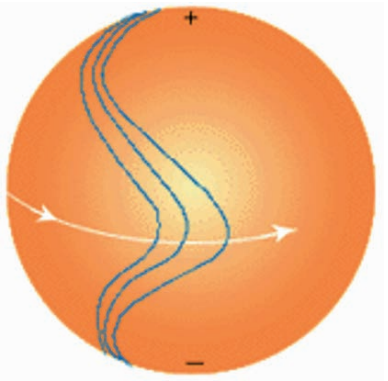


Figure 4.4: Illustration of the ω effect. The field lines are wrapped around because of the differential rotation of the Sun

The α Effect

However, the field lines are not only wrapped around the Sun but also twisted by the Sun's rotation. This effect is caused by the coriolis force. Because the field lines become twisted loops, this effect was called α effect. Early models of the dynamo assumed that the twisting is produced by the effects of the Sun's rotation on very large convective flows that transport heat to the Sun's surface. The main problem of that assumption was, that the expected twisting is too much and would produce magnetic cycles of only a couple of years. More recent dynamo models assume that the twisting is due to the effect of the Sun's rotation on rising flux tubes. These flux tubes are produced deep within the Sun.

The Interface between Radiation Zone and Convection Zone

If dynamo activity occurs throughout the entire convection zone the magnetic fields within that zone would rapidly rise to the surface and would not have enough time to experience either the alpha or the omega effect. This can be explained as follows: a magnetic field exerts a pressure on its surroundings ($\sim B^2$, proportional to its strength). Therefore, regions of magnetic fields will push aside the surrounding gas. This produces a bubble that rises continuously to the surface. However such a buoyancy is not produced in the radiation zone below the convection zone. Here, the magnetic bubble would rise only a short distance before it would find itself as dense as its surroundings. Consequently, it is assumed that magnetic fields are produced at this interface layer between the radiation zone and the convection zone.

Helioseismology has established the existence of a layer of strong gradients of angular velocity at the base of the solar convection zone. This layer, having a thickness of about 0.019 $R_\odot$, the *tachocline*, separates the convection zone exhibiting a strong latitudinal differential rotation from the radiative interior that rotates almost rigidly. Turbulence generated in the tachocline is likely to mix material in

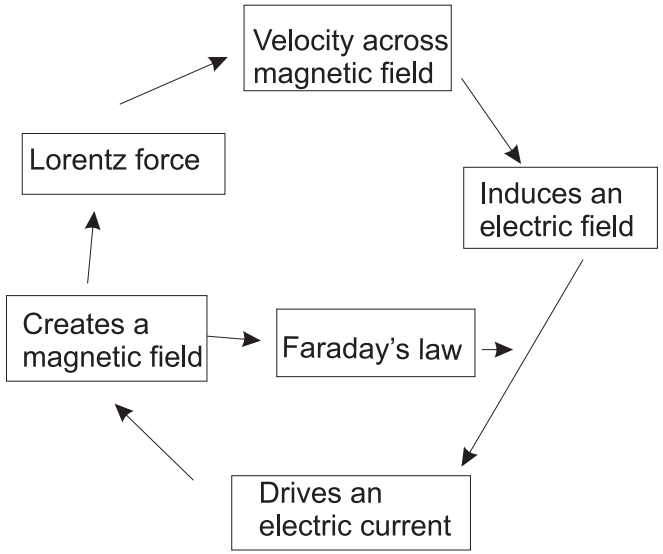


Figure 4.5: The MHD relation between flows and magnetic fields

the upper radiative zone resulting in the observed deficit of Li and Be. Gilman, 2005 [112] wrote a summary about the tachoclyne stressing its importance for in situ generation of poloidal fields as well as creating magnetic patterns that are seen on the surface.

The Meridional Flow

The solar meridional flow is a flow of material along meridional lines from the equator toward the poles at the surface and from the poles to the equator deep inside. At the surface this flow is in the order of 20 m/s, but the return flow toward the equator deep inside the Sun must be much slower since the density is much higher there- maybe between 1 and 2 m/s. This slow plasma flow carries material from the polar region to the equator in about 20 years.

Thus the energy that drives the solar dynamo comes from a) rotational kinetic energy, b) another part in the form of small-scale, turbulent fluid motions, pervading the outer 30% in radius of the solar interior (the convection zone).

4.2.3 Mathematical Description

Let us discuss some basic mathematics. In the magnetohydrodynamic limit the dynamo process is described by the induction equation:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \nabla \times (\eta_e \nabla \times \mathbf{B}) \tag{4.89}$$

The flow $\mathbf{u}$ is a turbulent flow. In the mean-field electrodynamics one makes the following assumptions: magnetic and flow fields are expressed in terms of a large-scale mean component and a small scale fluctuating (turbulent) component. If we average over a suitably chosen scale we obtain an equation that governs the evolution of the mean field. This is identical to the original induction equation but there appears a mean electromotive force term associated with the (averaged) correlation between the fluctuation velocity and magnetic field components. The basic principles of mean field electrodynamics were given by Krause and Rädler(1980) [174]. The velocity and the field are expressed as:

$$\mathbf{u} =< \mathbf{u} > +\mathbf{u}' \qquad \mathbf{B} =< \mathbf{B} > +\mathbf{B}' \tag{4.90}$$

$< \mathbf{u} >, < \mathbf{B} >$ represent slowly varying mean components and $\mathbf{u}', \mathbf{B}'$ non axisymmetric fluctuating components. The turbulent motion $\mathbf{u}'$ is assumed to have a correlation time τ and a correlation length λ which are small compared to the scale time t_0 and scale length l_0 of the variations of $< \mathbf{u} >$ and $< \mathbf{B} >$. In other words, τ is a mean time after which the correlation between $\mathbf{u}'(\mathbf{t} = \tau)$ and $\mathbf{u}'(\mathbf{t} = \mathbf{0})$ is zero and λ is comparable to the mean eddy size. We assume that $< \mathbf{u}' >, < \mathbf{B}' >= \mathbf{0}$.

This is substituted into the induction equation and subtracted from the complete equation:

$$\frac{\partial \mathbf{B}'}{\partial t} = \nabla \times (< \mathbf{u} > \times \mathbf{B}' + \mathbf{u}' \times < \mathbf{B} > +\mathbf{G}) - \nabla \times (\eta \nabla \times \mathbf{B}') \tag{4.91}$$

where

$$\mathbf{E} =< \mathbf{u}' \times \mathbf{B}' > \qquad \mathbf{G} = \mathbf{u}' \times \mathbf{B}' - < \mathbf{u}' \times \mathbf{B}' > \tag{4.92}$$

$\mathbf{E}$ is a mean electric field that arises from the interaction of the turbulent motion and the magnetic field. This field must be determined by solving the equation for $\mathbf{B}'$ and here several assumptions are made. First of all we stressed that $< \mathbf{v}' >= \mathbf{0}$. This may be a good assumption when considering a fully turbulent velocity field. However in the Sun we are dealing with a sufficiently ordered convective field where the Coriolis force plays an important role. The other approximation is a first order smoothing: $\mathbf{G} \sim \mathbf{0}$. That is valid only if $\mathbf{B}' << < \mathbf{B} >$. Then our equation reduces to:

$$\frac{\partial \mathbf{B}'}{\partial t} + \nabla \times (\eta \nabla \times \mathbf{B}') = \nabla \times (\mathbf{u}' \times < \mathbf{B} >) \tag{4.93}$$

We want to determine $\mathbf{E}$. Thus only $\mathbf{B}''$ the component of $\mathbf{B}'$ which is correlated with $\mathbf{u}'$ must be considered. By definition τ, $\mathbf{B}(\mathbf{t} + \tau)$ is not correlated with $\mathbf{B}(\mathbf{t})$ for any t. $\mathbf{B}''(\mathbf{t})$ may be determined by integration of the above equation from $t - \tau$ to t. Note also, that the order of the convective turn over time $\tau \sim \lambda / v$ and thus both $\mathbf{u}'$ and $< \mathbf{B} >$ may be regarded as independent of t. Thus the integration yields:

$$E_i = \alpha_{ij} < B_{ij} > +\beta_{ijk} \frac{\partial < B_j >}{\partial x_k} \tag{4.94}$$

where α_{ij}, β_{ijk} depend on the local structure of the velocity field and on τ. If the turbulent field is isotropic, then $\alpha_{ij} = \alpha\delta_{ij}$, $\beta_{ij} = \beta\epsilon_{ijk}$, and

$$\mathbf{E} = \alpha < \mathbf{B} > -\beta\nabla\times < \mathbf{B} > \tag{4.95}$$

If τ is small compared to the decay time λ^2/η, the diffusive term may be neglected and from 4.93 we get

$$\alpha = -\frac{1}{3}\tau < \mathbf{u}'.\nabla \times \mathbf{u}' >, \qquad \beta = \frac{\mathbf{1}}{\mathbf{3}}\tau\mathbf{v^2} \tag{4.96}$$

And finally:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\alpha\mathbf{B} + \mathbf{u} \times \mathbf{B}) - \nabla \times [(\eta + \beta)\nabla \times \mathbf{B}] \tag{4.97}$$

Compared to the normal induction equation, this contains the term $\alpha\mathbf{B}$ and the eddy-diffusivity coefficient β. In the mean field dynamo, the magnetic diffusivity η is replaced by a total diffusivity $\eta' = \eta + \beta$ and the equation becomes:

$$\frac{\partial B}{\partial t} = \nabla \times (\alpha\mathbf{B} + \mathbf{u} \times \mathbf{B}) + \eta'\nabla^{\mathbf{2}}\mathbf{B} \tag{4.98}$$

Please note that most often the prime is dropped on η; however, in the presence of α it is implied to use the turbulent diffusivity. It is assumed that $\mathbf{B}$ is axisymmetric. Then it can be represented by its poloidal and toroidal components $A(x, z, t)$ and $B(x, z, t)$ and $\mathbf{u} = \mathbf{u}(\mathbf{x}, \mathbf{z}, \mathbf{t})\mathbf{j}$. Neglecting the advection terms:

$$\left(\frac{\partial}{\partial t} - \eta\nabla^2\right) B = [\nabla u \times \nabla A].\mathbf{j} - \alpha\nabla^{\mathbf{2}}\mathbf{A} \tag{4.99}$$

$$\left(\frac{\partial}{\partial t} - \eta\nabla^2\right) A = \alpha B \tag{4.100}$$

Note that the dynamo action is possible because we have a regeneration of both toroidal and poloidal fields. Let us consider the source term in the first of the two above equations. ∇u describes a non uniform rotation. It can be argued that this term is larger than the next term involving α. This set of equations then describes the so called $\alpha - \omega$-dynamo. The equations describe:

- ω effect: the poloidal field is sheared by non uniform rotation to generate the toroidal field.
- α effect: this is the essential feedback. The helicity $\mathbf{v_c}.\nabla \times \mathbf{v_c}$ of the non axisymmetric cyclonic convection generates an azimuthal electromotive force $\mathbf{E}$ which is proportional to the helicity and to B_ϕ.

Let us define a characteristic length scale l_0, a decay time $t_0 = l_0^2/\eta$ and $u = s_0\omega$, where s_0 is of the order of the local radius of rotation and ω the local angular velocity. We may rewrite the above equations in terms of the non dimensional variables $t' = t/t_0$ and $\mathbf{r}' = \mathbf{r}/\mathbf{l_0}$. By an elimination of B and neglecting the α^2 terms we arrive at

$$\left(\frac{\partial}{\partial t'} - \nabla'^2\right)^2 A = \frac{\alpha l_0^2 s_0}{\eta^2}[\nabla'\omega \times \nabla' A].\mathbf{j} \tag{4.101}$$

If α_0 and ω_0 are scale factors giving the orders of magnitudes of α and $|\nabla' \omega|$ then

$$\left(\frac{\partial}{\partial t'} - \nabla'^2\right)^2 A = D\frac{\alpha}{\alpha_0}\left[\frac{\nabla'\omega}{\omega_0} \times \nabla' A\right].\mathbf{j} \tag{4.102}$$

In that equation the non dimensional *dynamo number* D is

$$D = \frac{\alpha\omega_0 l_0^2 s_0}{2\eta^2} \tag{4.103}$$

It is extremely important to note that the onset of a dynamo action depends on D. If D for a given system exceeds some critical value than there will be dynamo action. Examining our set of equations we may also note that dynamo action is possible when ∇u is negligible compared to α. Such dynamos are called α^2 dynamos. If both terms of the source term are comparable then we speak of an $\alpha^2\omega$ dynamo.

Solar like stars have well developed and structured convection zones. Thus, the $\alpha - \omega$ dynamo is the most likely dynamo mode.

Reviews on the solar dynamo and the emergence of magnetic flux at the surface can be found in Ossendrijver, 2003 [239], Fisher *et al.* (2000) [97] and Moreno-Insertis (1994) [226].

So far we have discussed large dynamos which are invoked to explain the origin of the solar cycle and of the large scale component of the solar magnetic field. We should add here that the origin of small scale magnetic fields can also be understood in terms of dynamo processes. Recent advances in the theory of dynamo operating in fluids with high electrical conductivity – fast dynamos, indicate that most sufficiently complicated chaotic flows should act as dynamos (Cattaneo, 1999 [57]). The existence of a large scale dynamo is related to the breaking of symmetries in the underlying field of turbulence (Cattaneo, 1997 [56]).

Steiner and Ferriz-Mas, 2005 [300], showed how solar radiance variability might be connected to a deeply seated flux-tube dynamo.

Observations form SOHO

Near the base of the convection zone the analysis of solar oscillations (data from the SOHO/MDI) has shown that there exist variations in the rotation rate of the Sun. A successive acceleration and deceleration with a strange period of 1.3 years was found near the equator and 1.0 years at high latitudes. The largest temporal changes were found both above and below the 'tachocline', a layer of intense rotational shear at the interface between the convection zone and the radiation zone (see Spiegel and Zahn, 1992 [297]). The variations near the equator are strikingly out of phase above and below the tachocline, and involve changes in rotation rate of about 6 nHz, which is a substantial fraction of the 30 nHz difference in angular velocity with radius across the tachocline. The solar magnetic dynamo is thought to operate within the tachocline, with the differential rotation there having a crucial role in generating the strong magnetic fields involved in the cycles of solar activity. This is illustrated in Fig. 4.6.

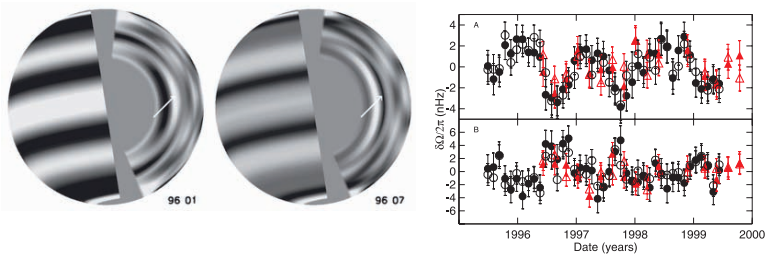


Figure 4.6: a) Cutaway images of solar rotation showing a peak and a trough of the 0.72R variation, with black indicating slow rotation, grey intermediate, and white fast. b) Variations with time of the difference of the rotation rate from the temporal mean at two radii deep within the Sun, with the site at 0.72 $R_\odot$ located above the tachocline and that at 0.63 $R_\odot$ below it, both sampling speeding up and slowing down in the equatorial region. Results obtained from GONG data for two different inversions are shown with black symbols, those from MDI with red symbols. (Image courtesy NSF's National Solar Observatory)

4.2.4 Solar Activity Prediction

Generally, prediction of solar activity is related to the problem of prediction of a given time series since solar activity parameters such as sunspot numbers are given as a function of time. Therefore, the problem can be examined on the basis of recent nonlinear dynamics theories. The solar cycle is very difficult to predict due to the intrinsic complexity of the related time behavior and to the lack of a successful quantitative theoretical model of the Sun's magnetic cycle. Sello (2001) [278] checked the reliability and accuracy of a forecasting model based on concepts of nonlinear dynamical systems applied to experimental time series, such as embedding phase space, Lyapunov spectrum, chaotic behavior. The model is based on a local hypothesis of the behavior on embedding space, utilizing an optimal number of neighbor vectors to predict the future evolution. The main task is to set up and to compare a promising numerical nonlinear prediction technique, essentially based on an inverse problem, with the most accurate prediction methods, like the so-called "precursor methods" which appear now reasonably accurate in predicting "long-term" Sun activity, with particular reference to "solar" and "geomagnetic" precursor methods based on a solar dynamo theory.

Snodgrass (2001) [289] studied azimuthal wind bands known as the *torsional oscillations* . These have been revealed primarily by studying the longitudinally averaged solar rotation over a period spanning several full solar rotations. This averaging yields what look like broad but slow, oppositely-moving ($\sim$5 m/s) bands lying to either side of the centroid of the sunspot butterfly, making the activity band appear to be a zone of weakly enhanced shear. The torsional pattern tells us something about the cycle, and since it precedes the onset of activity, it might be useful as a predictor of the level of activity to come. For cycle 23, the torsional pattern did not emerge until just before solar minimum, whereas for cycles 21 and 22 it appeared several years earlier. This would have suggested by 1996 the cycle 23 would be weaker than the previous two.

Calvo *et al.* (1995) [53] used the neural network technique to analyze the time series of solar activity (given by the Wolf number).

Hernandez (1993) [132] also used neural nets to construct nonlinear models to forecast the AL index (auroral electrojet index) given solar wind and interplanetary magnetic field (IMF) data.

Gleisner and Lundstedt (2001) [113] used a neural network-based model for prediction of local geomagnetic disturbances. Boberg *et al.* (2000) [38] made real time Kp predictions from solar wind data using neural networks.

4.3 Stellar Activity

4.3.1 Detection and Observation of Stellar Activity

The Sun is the only star that permits a two-dimensional study of its activity[5]. However, it is only a single set of stellar parameters, since its mass, composition and evolutionary status are fixed[6] . Stars are one-dimensional objects when observed from the Earth but they cover a wide range of physical parameters. Thus the solar-stellar connection is essential for a better understanding of solar phenomena as well as for stellar phenomena. In the 40s e.g. the solar chromosphere was thought to be unique.

Spectral line indicators for stellar activity are:

- EUV lines,
- Hα, He λ1083.0 nm,
- H and K lines of Ca II,
- Mg II.

The first detection of stellar activity phenomena was made by the observation of magnetic fields. Field strengths in the range of 1-2 kGauss can only be measured by a comparison of magnetically sensitive lines with magnetically insensitive lines. It was surprising that these stars seem to be covered by such strong fields about 20-80% of the total surface (the Sun is only covered $\leq$1%). The problem is, that by these methods coverages lower than 20% cannot be detected. That means that the Sun's magnetic field would not have been detected if it were at the distance of these stars. More than 100 years ago Pickering suggested that luminosity fluctuations in stars of the order of 20% over periods of days or a few weeks might indicate that they are spotted. In the 1970 extensive investigations were performed to explain luminosity variations of e.g. the RSCVn stars or BY Draconis stars (having luminosities $< 1/2\, L_\odot$). The observed lightcurves required circular spots. The RSCVn stars occur in binary stars were tidal interactions play an important role,

[5] A good textbook is C. Schrijver, Solar and Stellar Magnetic Activity Cambridge Astrophysics, 2000

[6] See also the new edition of Solar and Stellar Activity Cycles, P. R. Wilson, A. King, Cambridge 2004

therefore their starspots are quite different from the sunspots. BY Dra stars are rapidly rotating young low massive stars characterized by intense chromospheric emission. Large spots on the Sun cause a variation of the integrated flux $< 1\%$, whereas up to 30 % for RSCVn and BY Dra stars.

The size and extent of chromospheric active regions varies dramatically over the course of the activity cycle. Thus by measuring the H and K lines of other stars we can infer on stellar activity cycles.

How can we measure stellar parameters like differential rotation that play a key role in the onset of stellar dynamos? Let us assume we have a rapidly rotating spotted cool star and that it is observed one week apart. By comparing brightness/magnetic images of that star over such time intervals one can measure the rotation rates of starspots at different latitudes over several rotation cycles (Barnes *et al.* 2001) [26].

Also flares were detected on stars. Here it is extremely important to have observations in the EUV/X ray window. Generally pre main sequence stars show high levels of magnetic activity and strong flares. FU Orionis stars may be in a phase between T Tauri and post T Tauri stars. More details about that topic can be found in the review of Haisch *et al.* (1991) [124]. So far we have considered only stars which have an activity level by orders of magnitude larger than the Sun.

4.3.2 Stellar Activity Cycles

One of the programs that is being carried out since a long time is the HK project where the H and K activity of a large sample of stars is recorded. Almost 100 stars have been observed continuously since 1966; at present the project is monitoring long-term changes in chromospheric activity for approximately 400 dwarf and giant stars. In order to compare the data with the Sun, observations of reflected sunlight from the Moon are done at Mt. Wilson and at Sac Peak and Kitt Peak National Observatory. The sampling of the stars occurs rapidly: usually less than 10 min per star. The accuracy of the instrument is between 1% and 2%. When plotting the HK index against the $B - V$ color index (which is a measure for temperature as explained in chapter 1) then a clear trend can be seen. The HK index increases as the stellar temperature decreases. At this point one must be careful with the interpretation. It is not meant an absolute increase but a relative increase because in cooler stars also the continuum decreases.

In 1972 Skumanich [288] stated the $t^{-1/2}$ law for the time of stellar rotation and stellar chromospheric decay; the rotational velocity and the strength of the CaII emission of a late type star vary inversely with the square root of the star's age - *Skumanich law.* However later it was found that except massive T Tauri stars the majority of low mass stars rotates slowly.

It was also found that there exists a granulation boundary in the HRD at F5 III. Stars of later spectral type begin to develop a convective envelope that grows for the rest of their evolution. At the boundary these envelopes are extremely thin (only 3% of the star's radius). Stars on the right hand side in the HRD of the granulation border have smaller rotation rates.

In hydrodynamics, by definition, the *Rossby Number* is a ratio of inertial forces to the Coriolis Force for a rotating fluid. In astrophysics it is the ratio of the rotation period to the turnover time of the largest convective eddy. In stars with low Rossby numbers the rotation rate dominates the convective turnover time. The low Rossby number correlates well with the strong MgII 1940 emission. A low value of the Rossby number indicates a greater influence of the Coriolis forces. That means that the α effect becomes more important.

Stars can only be observed as point sources since we have no spatial resolution. Some stars show two simultaneous cycle periods. Other stars either have variable activity, or long trends in activity - longer than our 30-year baseline, or appear to be very inactive. For further details on that topics the book of Schrijver and Zwaan (2000) [274] is recommended where you find further references.

Dravins *et al.* (1993a) [79] made a detailed comparison of the current Sun (G2 V) with the very old solar-type star Beta Hyi (G2 IV) in order to study the post main-sequence evolution of stellar activity and of non thermal processes in solar-type atmospheres. This star has an age of 9.5 +/- 0.8 Gyr. The relatively high lithium abundance may be a signature of the early sub giant stage, when lithium that once diffused to beneath the main-sequence convection zone is dredged up to the surface as the convection zone deepens. Numerical simulations of the 3D photospheric hydrodynamics show typical granules to be significantly larger (a factor of about 5) than solar ones. The emission of the Ca H and K profiles was found to be weaker than that of the Sun. The observations suggest continuous changes in the chromospheric structure, rather than the sudden emergence of growth of active regions (Dravins *et al.*, 1993b [80])

Since several extrasolar planets have been found one should rise the question whether some of them might be suitable for life. Climatic constraints on planetary habitability were investigated by Kasting (1997) [153]. They found such zones around main sequence stars with spectral types in the early F to the mid K-range. The large amount of UV radiation emitted by early type stars poses a problem for evolving life in their vicinity. But there is also a problem with late-type stars; they emit less radiation at wavelengths < 200 nm which is required to split O_2 and initiate ozone formation. The authors show that Earth-like planets orbiting F and K stars may well receive less harmful UV radiation at their surfaces than does the Earth itself.

Chapter 5

The Earth's Atmosphere and Climate

In this chapter we give an overview of the Earth's atmosphere[1] and climate. We describe the possible evolution of the atmosphere and the variation of climate in the past (paleoclimatology). The influence of the Sun and its variation will be described in the next chapter.

5.1 The Earth's Atmosphere

5.1.1 Structure of the Atmosphere

The Earth's atmosphere is essential to life. It insulates the inhabitants of Earth from the extreme temperatures of space, filters out most radiation dangerous to life etc. It also provides the pressure that is necessary for liquid water at moderate temperatures on the surface. Considering the average temperature profile for the Earth's atmosphere, we can define the following regions:

- Troposphere: characterized by convective motions; warmer air is comparatively light and tends to rise, colder air is dense and tends to sink; the temperature decreases down to 200 K at it's upper boundary, the tropopause, at a height of 17 km. Most of the clouds and weather systems are located in the troposphere.

- Stratosphere: here the temperature slightly increases up to the stratopause at a height of about 50 km. In this layer there are no vertical motions, only horizontal motions occur. If a blub of air tends to rise it immediately becomes colder and thus denser and the buoyancy stops such motions. The temperature in this region increases gradually to -3^0 Celsius, due to the absorption of ultraviolet radiation by the ozone layer. 99% of "air" can be

[1] A good recent textbook is The Atmosphere, Frederick K. Lutgens, Edward J. Tarbuck, Dennis Tasa, Prentice Hall, 2003

Table 5.1: Composition of the Earth's atmosphere

Gas	Molecular weight	fraction by volume $[10^{-6}]$
N_2	28.02	780900
O_2	32.00	209500
Ar	39.94	9300
CO_2	44.10	300
CO	28.01	0.1
CH_4	16.05	1.52
N_2O	44.02	0.5
H_2O	18.02	$10^4...10^3$

found in the troposphere and stratosphere. The stratopause separates the stratosphere from the next layer.

- Mesosphere: the temperature decreases to -93^0 Celsius up to the mesopause (height 80 km). The mesopause is the coldest region in the atmosphere.

- Thermosphere (also called ionosphere): the temperature rises up to 1 000 K at a height of 250 km. In this region, thermal conduction becomes important. The extension is up to 600 km. The structure of the ionosphere is strongly influenced by the charged particle wind from the Sun (solar wind), which is in turn governed by the level of solar activity. A measurement of the structure of free electron density is an indicator for the degree of ionization.

Tropopause and troposphere are known as the lower atmosphere, stratosphere, stratopause, mesosphere and mesopause are called middle atmosphere and the thermosphere is called the upper atmosphere.

Incoming solar radiation with wavelength larger than 300 nm (in the visible part of the spectrum) penetrates down to the bottom. Radiation with $200\,\text{nm} < \lambda < 300\,\text{nm}$ is absorbed in the stratosphere (ozone layer) and solar radiation below 100 nm at higher layers.

Up to a height of about 100 km the composition is more or less constant. This is because of the high frequency of collisions between the molecules. These collisions become less efficient at heights above 100 km. The molecules experience a force of gravity that is proportional to their mass. Heavy gases are bound more closely to the Earth and lighter gases flout freely. Hence the lighter atomic oxygen is more abundant at heights above 160 km than the heavier nitrogen N_2. The region below 100 km is called *homosphere*, the region above 100 km the *heterosphere*.

Sunlight is absorbed in the atmosphere and this process is mainly responsible for its thermal structure. More than 50% of the energy incident from the Sun is absorbed by the surface. 30% is reflected back into space (20% from the clouds, 6% by air and 4% by the surface itself). The atmosphere absorbs only 16% of the incident solar energy. Most of this absorbed energy is captured by dust particles in the troposphere. If we want to construct a model of the atmosphere we have to take into account that it is exposed to two different radiation fields: a) from

Table 5.2: Energy received from the Sun at 1 AU

Wavelength range	Energy [ergcm^{-2}s^{-1} at 1 AU]
Solar constant	1.4×10^{6}
Solar wind	1
magnetic field	10^{-2}
4μm to ∞	700
300 nm to 4 μm	98% of total
120 nm to 300 nm	1.6×10^{4}
Lyman α	3-6
30 to 120 nm	2
3 to 30 nm	1
1 to 3 nm	0.01
0.01 to 1 nm	$10^{-3}...10^{-5}$
0 to 0.01 nm	$0..10^{-6}$

the Sun (covering all wavelengths from far UV to IR), and b) from IR radiation reflected at the surface of the Earth.

In Table 5.2 we give the energy received from the Sun at a distance of 1 AU in different wavelength regions. It is clearly seen, that 98% of the radiation from the Sun is in visible to the near IR.

The overall heat budget of the atmosphere is as follows: the surface receives 17% of its heat directly from the Sun, 15% from solar radiation scattered by clouds and 68 % from absorption of infrared radiation emitted by the atmosphere. What happens to the energy that is absorbed by the surface? The greater part (79%) is returned to the atmosphere in the form of radiation. The remainder part (21%) is transmitted to the atmosphere by conduction and as by product by the exchange of water H_2O. The surface cools when water evaporates and heat is transmitted to the air as vapor which recondenses to form clouds. Such phase transitions of H_2O play a major role in the energy budget of the lower atmosphere.

5.1.2 Composition

The composition of the Earth's atmosphere is given in Table 5.1.

Of course there are gases that can vary considerably both in space and time like nitric oxide, carbon monoxide and ozone. We can also consider the atmosphere as an extension of the biosphere, especially for gases like O_2, CO_2, CH_4, H_2. Oxygen is produced by photosynthesis:

$$h\nu + CO_2 + H_2O \longrightarrow CH_2O + O_2 \tag{5.1}$$

In this formula CH_2O denotes any variety of organic compounds. Aerobic respiration and decay occur in the reverse reaction:

$$CH_2O + O_2 \longrightarrow CO_2 + H_2O + \text{energy} \tag{5.2}$$

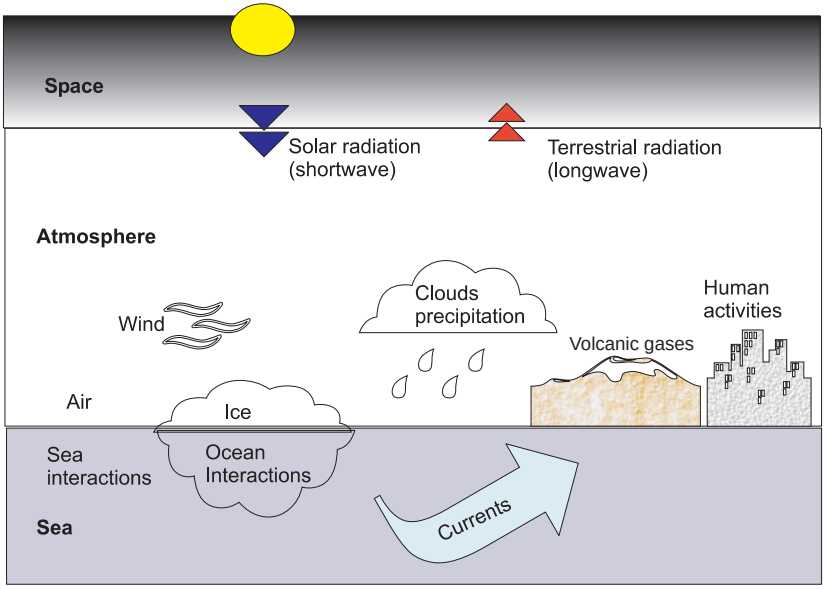


Figure 5.1: Major elements of the climate system

In the absence of this reaction, carbon would accumulate in organic form and the fuel for photosynthesis (atmospheric CO_2) would be depleted. If the supply of O_2 is limited such as in the sediments of organic rich swamps and in the stomachs of ruminants, we get as a product methane CH_4.

In Table 5.3 the change of the greenhouse gas and other gas concentrations of the Earth's atmosphere is given.

In Table 5.3 the present tropospheric concentration estimates are calculated as annual arithmetic averages; ppm = parts per million (10^6), ppb = parts per billion (10^9), ppt = parts per trillion (10^{12}).

The Global Warming Potential (GWP) is generally used to contrast different greenhouse gases relative to CO_2. The GWP provides a simple measure of the relative radiative effects of the emissions of various greenhouse gases and is calculated using the formula:

$$GWP = \frac{\int_0^n a_i c_i dt}{\int_0^n a_{CO_2} c_{CO_2} dt} \tag{5.3}$$

where a_i is the instantaneous radiative forcing due to a unit increase in the concentration of trace gas, i, c_i is concentration of the trace gas, i, remaining at time, t, after its release and n is the number of years over which the calculation is performed. This formula is taken from the Intergovernmental Panel on Climate Change (IPCC) (Houghton *et al.*, 1990 [138]).

How was the concentration of greenhouse gases in the past? Ice cores, cylinders of ice drilled out of glaciers and polar ice sheets, have played an important role in to answer such questions. The drilling for The Greenland Ice Sheet Project

Table 5.3: Current Greenhouse Gas Concentrations and Other Components

Gas	Pre-industrial conc. (1860)	Present tropospheric conc.	GWP 100 yr time horizon	Atm. lifetime (yr)
CO_2 (ppm)	288	369.4[1]	1	120
CH_4 (ppb)	848[2]	1839[3]/ 1726[4]	23	12
N_2O (ppb)	285[5]	315[3]/ 314[4]	296	114
CCl_3F (ppt)	zero	263[3]/ 260[4]	3800	50
CF_2Cl_2 (ppt)	zero	544[3]/ 537[4]	8100	102
$C_2F_3Cl_3$ (ppt)	zero	82[3]/ 82[5]	4800	85
surface ozone (ppb)	2517	2418/ 2919	20	hours

The measurements are from: [1] in situ air samples collected at Mauna Loa Observatory, Hawaii (Bacastow *et al.*1985 [19]). [2] Etheridge, D. M.; Pearman, G. I.; Fraser, P. J. , Tellus, Series B - Chemical and Physical Meteorology, 44B, no. 4, 282. These authors used an ice core from the antarctic called DE08. The extracted ice-core air is analyzed for methane using gas chromatography with flame-ionization detection. The mean air-age was 35 yr younger than the host ice. [3] Values from Macehead, Ireland. [4] Cape Grim, Tasmania [4] data from Law Dome BHD ice core, Etheridge *et al.*, 1988 [92]

Two began in 1989, more than 3000 m deep. In 1992 there were data available to reconstruct the climate over the past 200 000 years.

The CO_2 data are from an ice core analyzed by Neftel *et al.* (1985) [230]. An example of their measurements is given in Table 5.4.

These measurements of the CO_2 gas concentration enclosed in an ice core from Siple Station, Antarctica, indicate that atmospheric CO_2 concentration around 1750 was 280 ± 5 ppmv (parts per million per volume) and has increased since, essentially because of human factors, by 22.5 percent to 360 ppmv around 2000. The anthropogenic emission of CO_2 is about 7 Gt/yr. The natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air in Antarctic ice and firn was described in Etheridge *et al.* (1996) [93].

5.1.3 Paleoclimatology

First of all let us give a definition of climate: Climate is the weather we expect over the period of a month, a season, a decade, or a century. More technically, climate is defined as the weather conditions resulting from the mean state of the atmosphere-ocean-land system, often described in terms of "climate normals" or average weather conditions. Climate Change is a departure from the expected average weather or climate normals.

A better knowledge about the climate and its variations in the past will enable us to better understand what forces climate and its variations in the future. Since there exists only a 140 years instrumental record, we have to use proxies to reconstruct climate in the past. Some widely used proxy climate data are (see 5.2):

Table 5.4: Historical CO_2 record from the Siple Station Ice Core

Depth [m]	Samples measured	Date of ice	Date of Air enclosed	CO_2 conc. (ppmv) in extracted air
187.0-187.3	10	1663	1734-1756	279
177.0-177.3	10	1683	1754-1776	279
162.0-162.3	9	1723	1794-1819	280
147.0-147.2	10	1743	1814-1836	284
128.0-129.0	47	1782	1842-1864	288
111.0-112.0	26	1812	1883-1905	297
102.0-103.0	26	1832	1903-1925	300
92.0-93.0	25	1850	1921-1943	306
82.0-83.0	28	1867	1938-1960	311
76.2-76.6	11	1876	1947-1969	312
72.4-72.7	11	1883	1954-1976	318

- Historical data: Historical documents contain a wealth of information about past climates (diaries, records...).
- Corals: Corals build their hard skeletons from calcium carbonate, a mineral extracted from sea water. The carbonate contains oxygen and the isotopes of oxygen, as well as trace metals, that can be used to determine the temperature of the water in which the coral grew. These temperature recordings can then be used to reconstruct climate during that period of time when the coral lived. Increased sea surface temperature has negative effects on the health of coral. The most visible symptom of declining coral health is coral bleaching.
- Fossil pollen: Each species and genus of plants produces pollen grains which have a distinct shape. These shapes can be used to identify the type of plant from which they came. Pollen grains are well preserved in the sediment layers that form in the bottom of a pond, lake or ocean; an analysis of the pollen grains in each layer tells us what kinds of plants were growing at the time the sediment was deposited. Inferences can then be made about the climate based on the types of plants found in each layer.
- Tree rings: Since tree growth is influenced by climatic conditions, patterns in tree-ring widths, density, and isotopic composition reflect variations in climate. In temperate regions where there is a distinct growing season, trees generally produce one ring a year, and thus record the climatic conditions of each year. Trees can grow to be hundreds to thousands of years old and can contain annually-resolved records of climate for centuries to millennia.
- Ice cores: Located high in mountains and deep in polar ice caps, ice has accumulated from snowfall over many centuries. These cores contain dust, air

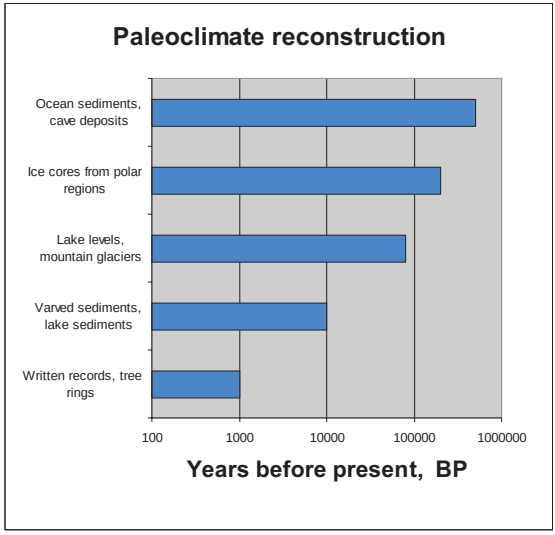


Figure 5.2: Timelines for various paleoclimate reconstruction methods.

bubbles, or isotopes of oxygen, that can be used to interpret the past climate of that area[2]. Let us briefly discuss one example of isotope measurements: Of the temperature dependent markers the most important is the ratio of ^{18}O to ^{16}O. This can be explained by the fact that water molecules composed of $H_2^{18}O$ evaporate less rapidly and condense more readily than water molecules composed of $H_2^{16}O$. Thus, in the ice cores one obtains annual layers starting with ^{18}O rich, becoming ^{18}O poor, and ending up ^{18}O rich.

- Volcanic eruption: After the eruption of volcanoes, the volcanic ash and chemicals are washed out of the atmosphere by precipitation and these eruptions leave a distinct marker within the snow which washed the atmosphere.

  We can then use recorded volcanic eruptions to calibrate the age of the ice-core (here the deuterium to hydrogen ratio is an important proxy).

  Ice cores from Vostok, Antarctica, were the first to cover a full glacial-interglacial cycle.

- Ocean and lake sediments: Between 6 and 11 billion tons of sediment (tiny fossils and chemicals) accumulate in the ocean and lake basins each year.

How can we infer e.g. from ice cores past climate? The accumulation which is governed by saturation water pressure was lower during colder periods and vice versa. Accumulation rates inferred in this way are supported by measurements of the cosmogenic isotope Beryllium 10 (^{10}Be), an isotope produced by the interaction

[2]see also the textbook of R.B. Alley, The Two-Mile Time Machine, Princeton Univ. Press, 2002

of cosmic rays and the upper atmosphere, can be used to determine past snow accumulation in Vostok ice. Deposition of ^{10}Be is assumed to be constant. The other two elements which are important are ^{18}O and deuterium. In Antarctica, a cooling of 1^0C results in a decrease of 9 per mil deuterium. The last ice age is characterized by three minima separated by slightly warmer episodes called interstadials.

Air initially enclosed in Vostok ice provides our only record of variations in the atmospheric concentrations of CO_2 and CH_4 over a complete glacial-interglacial cycle. For both greenhouse gases, concentrations were higher during interglacial periods than during full glacial periods.

Crowley (2000) [70] discussed the causes of climate change over the past 1000 years. His main conclusion is that as 41-64% of pre-anthropogenic (pre-1850) decadal-scale temperature variations were due to changes in solar irradiance and volcanism.

Several periods of warmth (listed below) have been hypothesized to have occurred in the past. However, upon close examination of these warm periods, it becomes apparent that these periods of warmth are not similar to 20th century warming for two specific reasons: a) the periods of hypothesized past warming do not appear to be global in extent, or b) the periods of warmth can be explained by known natural climatic forcing conditions that are uniquely different than those of the last 100 years.

Examples of periods of warmth:

- Medieval: $\sim$ 9th to 14th centuries; this seems to be in doubt now because the temperature anomaly at that time was very small; however the Little Ice Age for the northern hemisphere from 15th to 19th centuries is clearly seen (Fig. 5.3).
- mid-Holocene warm Period (approx. 6 000 years ago); this seems to be in connection with changes of the Earth's orbit (Theory of Milankovich).
- Penultimate interglacial period (approx. 125 000 years ago). It appears that temperatures (at least summer temperatures) were slightly warmer than today (by about 1 to 2^0C), caused again by the changes in the Earth's orbit (Hughes and Diaz, 1994 [142]).
- Mid-Cretaceous Period (approx. 120-90 million years ago): Breadfruit trees apparently grew as far north as Greenland (55° N), and in the oceans, warm water corals grew farther away from the equator in both hemispheres. The mid-Cretaceous was characterized by geography and an ocean circulation that was vastly different from today, as well as higher carbon dioxide levels (at least 2 to 4 times higher than today).

5.1.4 Theory of Milankovich

Seasons on Earth are caused by the tilt of the Earth's rotation axis relative to its plane of revolution around the Sun[3]. In summer, one hemisphere is pointing

[3] which is called the ecliptic

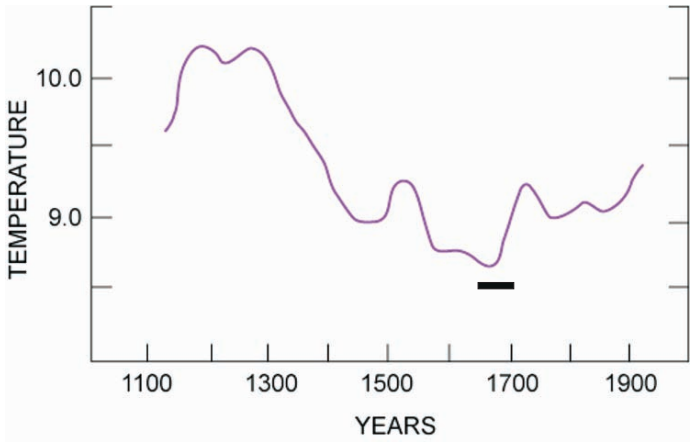


Figure 5.3: Temperature anomaly clearly showing the Little Ice Age. The Maunder minimum is marked. (redrawn from Lamb, 1977 [185])

toward the Sun, at the same time the opposite hemisphere is in winter. If the Earth's axis were not inclined every point on the earth would receive the same amount of sunlight each day of the year. Changes in this tilt can change the severity of the seasons. More tilt means more severe seasons, i.e. warmer summers and colder winters. The tilt of the Earth's axis changes between 22 and 25 degrees on a cycle of about 41 000 years. If the summers are cool snow and ice last from year to year in high latitudes building up massive ice sheets. Now positive feedbacks in the climate system start to work. Snow reflects more of the sun's energy into space causing additional cooling. Also the amount of the greenhouse gas CO_2 in the atmosphere falls as ice sheets grow and thus adding to the cooling.

Another astronomical effect on climate is that the orbit of the earth is not circular. Presently, perihelion (closest approach to the Sun) occurs in January, thus on the northern hemisphere winters are slightly milder. The perihelion changes in a cycle of 22 000 years. Therefore, 11 000 years ago perihelion occurred in July making seasons more severe than today. The eccentricity of the earth's orbit varies on cycles of 100 000 and 400 000 years. It is the combined effect of the 41 000 year tilt cycle and the 22 000 year perihelion cycle plus the small effect from the eccentricity that influences the climate. These variations of the Earth's orbit were first investigated by Milankovich.

To study the effect of these astronomical variations on climate one must take into account, that orbital changes occur over thousands of years and the climate system also takes thousand of years to respond. The primary driver of ice ages seems to be the total summer radiation received in northern latitude zones near 65^0 north (65N) (this is where the major ice sheets formed in the past) and past ice ages correlate with the 65N summer insolation. Astronomical calculations show that the 65N summer insulation should increase gradually over the next 25 000 years. No decline of the 65N summer insolation that is sufficient to cause an ice age is expected within the next 100 000 years.

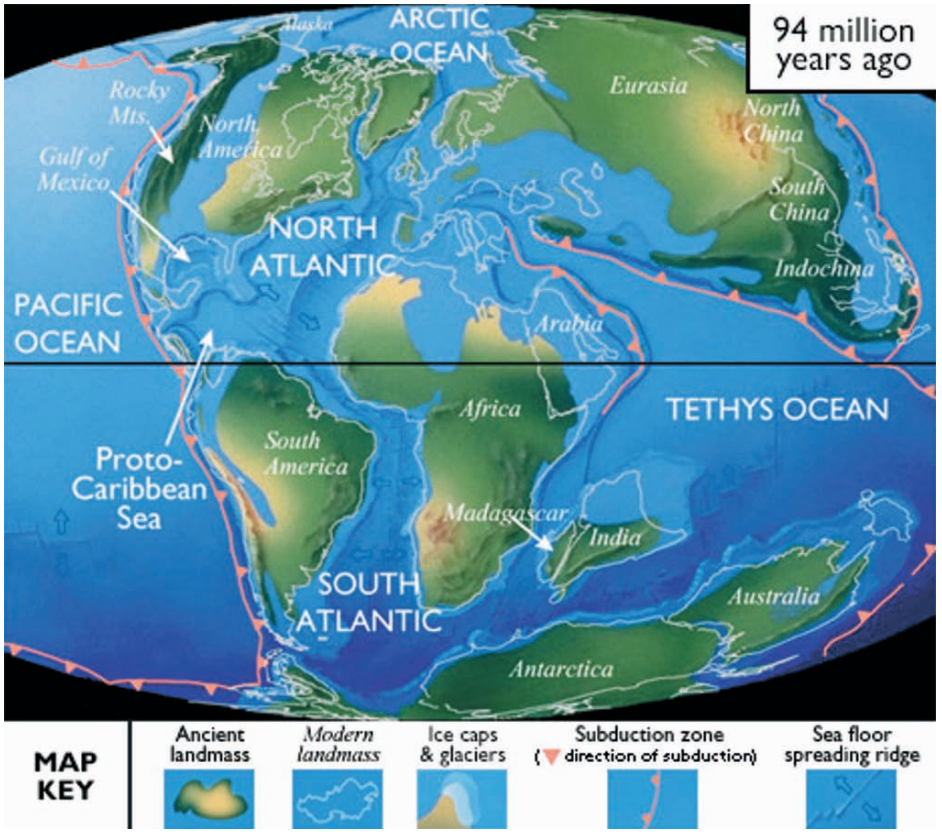


Figure 5.4: Cretaceous climate and land/sea distribution. During the Late Cretaceous the global climate was warmer than today's climate. No ice existed at the Poles. Dinosaurs migrated between the Warm Temperate and Cool Temperate Zones as the seasons changed. The sea level was about 100 m higher than today. Courtesy: *http* : *//www.paleoportal.org/timespace/*

Warm interstadials have always been accompanied by an increase of the atmospheric concentration of the three principal greenhouse gases. This increase has been, at least for CO_2, vital for the ending of glacial epochs. A highly simplified course of events for the past four transitions would then be as follows:

- changing orbital parameters initiated the end of the glacial epoch,
- an increase in greenhouse gases then amplified the weak orbital signal,
- in the second half of the transition, warming was further amplified by decreasing albedo, caused by melting of the large ice sheets in the Northern Hemisphere going parallel with a change of the ocean circulation.

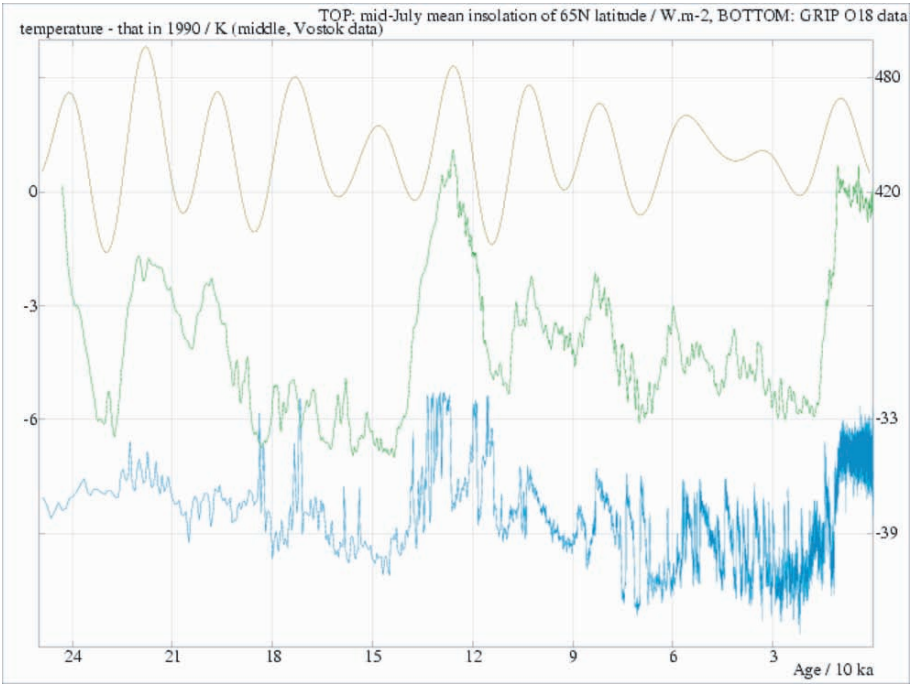


Figure 5.5: Upper curve: average insolation of 65 degrees northern latitude (Watts per one square meter of a horizontal atmosphere) in mid-July. As seen, it varies from some 390 to 490 W/m^2. Middle curve: Global temperature (Vostok ice core). Lower Curve: Greenland, GRIP core. Image courtesy: Jan Hollan

The isotopic records of Greenland ice cores show evidence for fast and drastic climatic changes during the last glacial epoch. Possible causes and mechanisms of such changes and their significance as global climatic events are discussed by Stauffer (2000) [299]. Ice core results also enable the reaction of the environment to past global changes to be investigated. The deglaciation of the northern hemisphere is described in Alley and Clark (1999 [5]). A carbon cycle model was used to reconstruct the global mean surface temperature during the last 150 Million years showing that during this period the tectonic forcing such as decrease in volcanic activity and the formation and uplift of the Himalayas and the Tibetan Plateau dominated the control of the climate (Tajika, 2001) [308].

5.1.5 Greenhouseffect

Trace constituents of the atmosphere such as H_2O, CO_2, O_3 absorb energy at longer wavelengths and thus trap heat radiated by the surface. The effect is very similar to that of a glass pane in a greenhouse. The atmosphere is transparent to solar radiation but it is opaque to longer wavelengths. The infrared absorbing gases return heat to the ground and account for about 70% of the net input of energy to

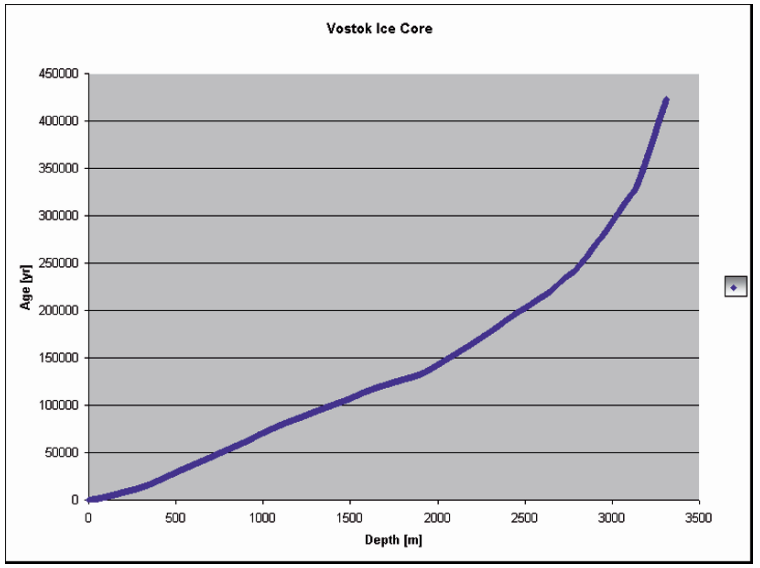


Figure 5.6: Vostok Ice core. Different depth can be attributed to different age

the surface. If our atmosphere would contain no water vapor and carbondioxide, the surface temperature would be about 40 K colder than today. This would imply that large portions of the planet would be covered with ice.

Since the 1980s there is a growing concern that the increase in the abundance of carbondioxide caused by combustion of fossil fuels could lead to a general warming of the global climate (see Fig. 5.8). Similar greenhouse effects arise from the gases methane, nitrous oxide and chlorofluorocarbons (CFCs). All these gases are referred to as greenhouse gases due to their ability to trap heat.

The variation of greenhouse gases was described before.

5.1.6 Ozone

The ozone is measured in Dobson units. 1 Dobson Unit (DU) is defined to be 0.01 mm thickness at STP (standard temperature and pressure) as the physical thickness would be if compressed in the Earth's atmosphere. The ozone layer is very thin a normal range is 300 to 500 Dobson units.

We can make the simplification that throughout the stratosphere all of the radiative energy from the sun that is absorbed by O_3 is converted locally to heat. The heating rate depends on the distribution of Ozone with height and on the incoming solar energy.

The absorption of shortwave solar radiation in altitudes above the troposphere is responsible for the temperature increase in these layers. Ozone absorbs most of the UV portion of sunlight ($200 < \lambda < 300\,\text{nm}$). The absorption process results in the dissociation of O_3.

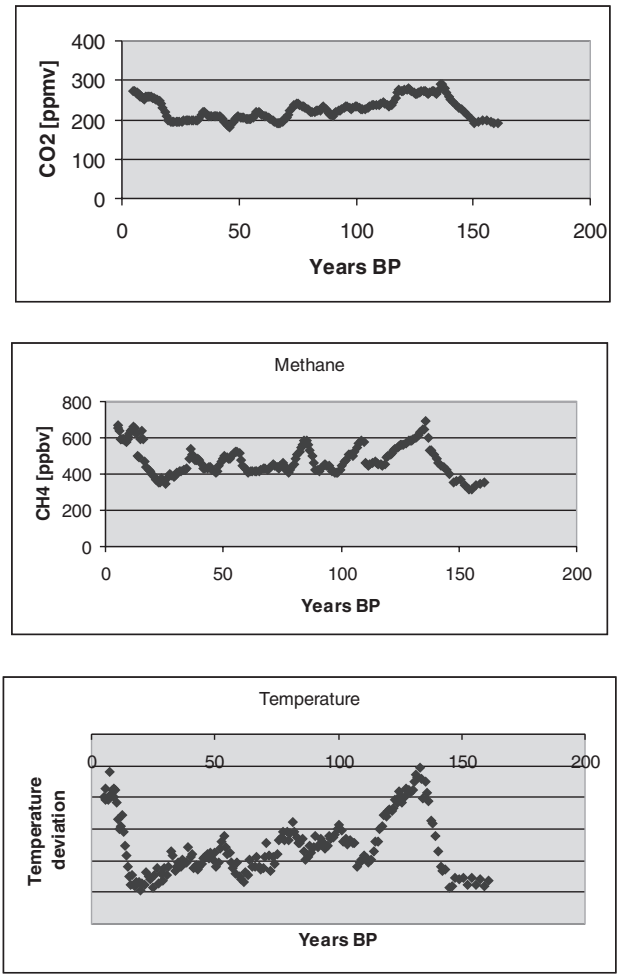


Figure 5.7: Variation CO_2 and CH_4 in parallel with temperature from Vostok climate records. Years BP are given in units of 1000. Credits: National Ice Core Laboratory.

The Chapman reactions describe the formation of ozone. First the following reaction leads to a photodissociation of oxygen:

$$O_2 + h\nu \rightarrow O + O \qquad \lambda < 175\,\text{nm} \tag{5.4}$$

Then a recombination of oxygen occurs:

1. by direct two body reaction (reaction very rare!)

$$O + O \rightarrow O_2 + h\nu \tag{5.5}$$

2. recombination by a three body process:

$$O + O + M \rightarrow O_2 + M \tag{5.6}$$

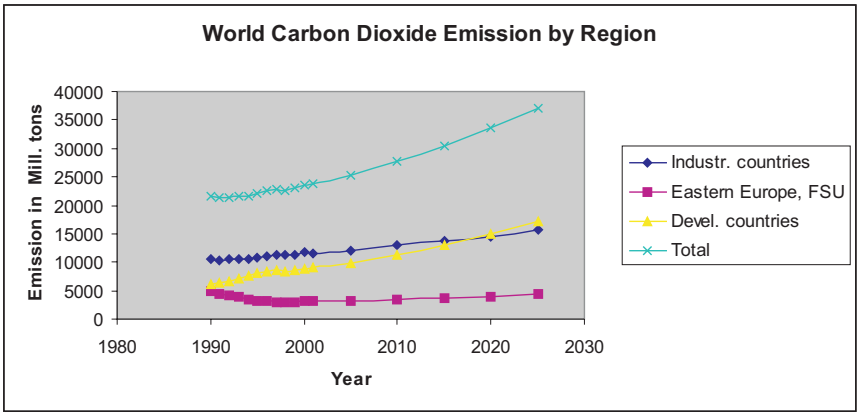


Figure 5.8: World Carbon Dioxide Emissions. It is seen that the CO_2 emission from the Eastern European Countries and the former Soviet Union (FSU) declines, e.g. further restructuring of the coal mining industries in Poland and the Czech Republic (US Dept. of Energy)

$$O + O_2 + M \rightarrow O_3 + M \tag{5.7}$$

The destruction of ozone occurs by the reactions: a) photodissociation

$$O_3 + h\nu \rightarrow O_2 + O \qquad \lambda < 310\,\text{nm} \tag{5.8}$$

and b) by the reaction:

$$O + O_3 \rightarrow O_2 + O_2 \tag{5.9}$$

Ozone peaks in number density at altitudes about 30 km.

It is now well known that ozone can be easily destroyed by several reactions. We just give a few examples:

Destruction of ozone by free hydrogen atoms

$$H + O_3 \rightarrow OH + O_2 \tag{5.10}$$

Free atomic H is produced from H_2O, CH_4.

Nitrogen oxides, chlorine and halomethanes act also as catalysts to destruct ozone. The problem here is, that they react with ozone, destroying it but remain unchanged. One example:

$$NO + O_3 \rightarrow NO_2 + O_2 \tag{5.11}$$

Natural events such as volcanic eruptions can strongly influence the amount of Ozone in the atmosphere. However, man-made chemicals such as CFCs or chlorofluorocarbons are now known to have a very dramatic influence on Ozone levels too. CFCs were once widely used in aerosol propellants, refrigerants, foams, and

industrial processes. Changes in the ozone layer caused by release of CFC's in the atmosphere have the potential of producing biological damage through increased UVB radiation[4]. While cloud cover provides protection on the ground against solar radiation in the visible and near UV wavelengths, biologically damaging radiation near 300 nm is controlled primarily by the total ozone content.

So far we have discussed ozone in the stratosphere. In the Earth's lower atmosphere, near ground level, ozone is formed when pollutants emitted by cars, power plants, industrial boilers, refineries, chemical plants, and other sources react chemically in the presence of sunlight. Ozone at ground level is a harmful pollutant. Ozone pollution is a concern during the summer months, when the weather conditions needed to form it, lots of sun, hot temperatures, normally occur. Tropospheric ozone is either produced by oxidation of hydrocarbons and CO or by downward transportation of stratospheric ozone. Some examples of reactions are given below.

$$\begin{aligned} \mathrm{CO} + \mathrm{OH} &\rightarrow \mathrm{CO_2} + \mathrm{H} \\ \mathrm{H} + \mathrm{O_2} + \mathrm{M} &\rightarrow \mathrm{HO_2} + \mathrm{M} \\ \mathrm{HO_2} + \mathrm{M} &\rightarrow \mathrm{OH} + \mathrm{NO_2} \\ h\nu + \mathrm{NO_2} &\rightarrow \mathrm{NO} + \mathrm{O} \\ \mathrm{O} + \mathrm{O_2} + \mathrm{M} &\rightarrow \mathrm{O_3} + \mathrm{M} \\ \mathrm{CO} + 2\mathrm{O_2} &\rightarrow \mathrm{CO_2} + \mathrm{O_3} \end{aligned}$$

5.1.7 The Structure of the Higher Atmosphere

Temperature Inversion in the Thermosphere

There is a wide range of textbooks covering that topic[5]. Above 80 km there is an inversion of the temperature that is caused by the absorption of solar radiation below 200 nm. Let us briefly discuss the most important processes:

- $100\,\mathrm{nm} < \lambda < 200\,\mathrm{nm}$: absorption of solar radiation leads to a dissociation of $\mathrm{O_2}$:

$$h\nu + \mathrm{O_2} \rightarrow \mathrm{O} + \mathrm{O} \tag{5.12}$$

- shorter wavelengths: ionization of $\mathrm{O}, \mathrm{O_2}, \mathrm{N_2}$:

$$\begin{aligned} h\nu + \mathrm{O} &\rightarrow \mathrm{O^+} + e \\ h\nu + \mathrm{O_2} &\rightarrow \mathrm{O_2^+} + e \\ h\nu + \mathrm{N_2} &\rightarrow \mathrm{N_2^+} + e \end{aligned}$$

[4]UV radiation is divided by wavelength into UVA (320-400nm), UVB (290-320nm) and UVC (100-290nm).

[5]Chemistry of the Upper and Lower Atmosphere, Barbara J. Finlayson-Pitts, James N. Pitts, Academic Press, 1999

The electrons that are emitted by these reactions loose energy by collision, elastic and inelastic. This can cause further ionization and contribute to the production of excited states and the associated emission of airglow.

Electrons can be removed by dissociative recombination:

$$\begin{aligned} e + O_2^+ &\rightarrow O + O \\ e + NO^+ &\rightarrow N + O \end{aligned}$$

There is of course a balance between dissociation of O_2 and recombination.

- Reformation of molecular oxygen:

$$\begin{aligned} O + O_2 &\rightarrow O_3 + M \\ H + O_3 &\rightarrow OH + O_2 \\ O + OH &\rightarrow O_2 + H \end{aligned}$$

The recombination of oxygen is catalyzed by the presence of hydrogen. Such catalytic reactions play an important role in the chemistry of the atmosphere below 80 km. The density in the thermosphere is low, therefore O diffuses downward. The recombination requires higher densities and is confined to regions below 100 km. The dissociation of O_2 can occur at any level.

Hydrogen Loss

Any particle in the atmosphere is bound to the Earth by the force of gravity. If we move such a particle a vertical distance Δz then the work $mg\Delta z$ is done. m denotes the mass of the particle, g the gravitational acceleration $=9.81\,\mathrm{m/s^2}$. The work that must be done to escape the gravitational field is mgR, where R is the radius of the Earth $\sim 6\,400\,\mathrm{km}$. All atoms or molecules have a range of speeds that is described by the Maxwell–Boltzmann distribution. The average kinetic energy is given by:

$$E_{\mathrm{kin}} = \frac{3}{2}kT \tag{5.13}$$

where k is the Boltzmann constant ($1.38 \times 10^{-16}\,\mathrm{erg\,K^{-1}}$). Thus an atom can escape the gravitational field if its thermal kinetic energy $\sim kT$ is much larger than mgR. Of course we must also consider collisions (except at the highest level in the atmosphere). At the high temperatures in the thermosphere (700...2 000 K), significant numbers of hydrogen atoms have velocities above the escape velocity $v_{\mathrm{esc}} \sim 11.2\,\mathrm{km/s}$. Therefore, hydrogen is lost at a rate of 10^8 atoms per $\mathrm{cm^2}$ per second. These escaping hydrogen atoms are derived mainly from the oceans and over the past 4.5×10^9 years of the Earth's history, the sea level has declined by two meters globally. Of course during this reaction also O_2 is set into the atmosphere which was crucial for the evolution of life.

There is also a significant loss of helium.

5.2 Earth's History and Origin of the Atmosphere

In this section we discuss the main steps in the evolution of the Earth and life on the Earth. Also the problem of the formation of the atmosphere will be shortly addressed.

5.2.1 History of the Earth

The history of Earth is recorded in the igneous, sedimentary and metamorphic rocks of the outer crust called the outer lithosphere as is sketched in Fig. 5.9. The formation of mountain ranges like Andes, Alps, Himalayas, and Rocky Mountains took place in the Cretaceous. During the Permian there was a widespread glaciation, mountains rising and the atmospheric CO_2 and O was reduced. In the Cambrian atmospheric oxygen reaches the first critical level, in the Silurian it reaches the second critical level. Stromatolithes are interpreted as calcareous algaes, which lived in the Precambrian and Paleozoic and to some degree up to the Triassic. They grew on beaches in layers, one upon the other, producing cauliflower-like forms.

The bluegreen algaes, nowadays also called cyanobacteria[6] were the "inventors" of the photosynthesis, they cannot be discriminated as animals or plants. They lived perhaps without rivalry on empty beaches, and the atmosphere contained only 0.2 % oxygen contrary to 20 % nowadays. With their production of O_2 they are the beginning of the rise of the oxygen-contents of the atmosphere. The prokaryotic[7] Cyanobacteria are both photosynthetic and aquatic living in water and producing their own food (autotrophic). They are unicellular bacteria but exist in great colonies and are the oldest known fossils (up 3.5×10^9 years old) but still they constitute one of the largest and most important group of bacteria on Earth. The oxygen atmosphere that we depend on was generated by numerous cyanobacteria during the Archaean and Proterozoic Eras. Before that time, the atmosphere had a very different chemistry. The other great contribution of the cyanobacteria is the origin of plants. The chloroplast with which plants make food for themselves is actually a cyanobacterium living within the plant's cells.

The plants first appeared in the Ordovician, but did not begin to resemble modern plants until the Late Silurian. By the close of the Devonian, about 360 million years ago, there was a wide variety of shapes and sizes of plants around, including tiny creeping plants and tall forest trees.

The most striking, and important, feature of plants is their green color, the result of a pigment called chlorophyll. Plants use chlorophyll to capture light energy, which fuels the manufacture of food–sugar, starch, and other carbohydrates. Without these food sources, most life on earth would be impossible.

The paleozoic era was the times of supercontinents. In the Cambrian there was a breakup of the global continent Rodinia, and at the end of the Paleozoic the formation of Pangea started as the Earth's continents came together once again.

[6] the Greek cyanos means blue

[7] Prokaryotes are single celled organisms that do not have a nucleus, mitochondria or any other membrane bound organelles.

History of the Earth				
Praecambrian	570 Mill. BP		One celled organisms, prokaryotes	
Paleozoic	570-240 Mill. BP			
	570-500	Cambrian	Multicellular life	
	500-435	Ordovician	Primitive life on land, vertebrates in ocean	
	435-410	Silurian	First plants, insects on land	
	410-360	Devonian	Spiders, mites, amphibians	
	360-290	Carboniferous	First true reptiles, coals begin to form	
	290-240	Permian	Mysterious mass extinction of life; 90 % of all organisms die out; reptiles inherit Earth	
Mesozoic	240-65 Mill. BP			
	240-205	Triassic	Small dinosaurs, ichtyosaurs, first true mammals	
	205-138	Jurassic	Huge dinosaurs, flying pterosaurs, oldest known birds	
	138-65	Cretaceous	Global warming, spread of dinosaurs. At the end sudden mass extinction (asteroid impact), 70 % of all organisms died	
Cenozoic	65 Mill. BP - present			
	65-1.6	Tertiary		
		65-55	Paleocene	Mammals inherit Earth
		55-38	Eocene	Ancestral forms of horses, rhinoceros, camel and others like bats, primates. Mammals adapt to marine life.
		38-24	Oligocene	Elephants, cats, dogs, monkeys
		24-5	Miocene	Global climate cools; establishment of the Antarctic ice sheet; large apes in Africa and southern Europe
		5-1.6	Pliocene	Climate becomes cooler and drier. Mammals dominant life form; ancestors of modern humans.
	1.6-present	Quaternaray		
		1.6 Mill-10000 y	Pleistocene	Most recent global ice age; glacier ice spreads out over more than 25 % of Earth's land surface; modern humans arise
		10000 y-present	Holocene	Global climate moderates; ice sheet retreat from Europe and North America; rise of sea levels

Figure 5.9: The history of the Earth. BP means before present

5.2.2 Origin of the Atmosphere

Let us start with the remark that the origin of our earth's atmosphere is still subject to much speculation. However the most probable history of its evolution was as follows[8].

Our Earth was formed some 4.5 billion years ago. At that time it was probably too hot to retain any primordial atmosphere. This first atmosphere most probably consisted of helium, hydrogen, ammonia and methane. At that time the Earth was

[8]see also: The Chemical Evolution of the Atmosphere and Oceans (Princeton Series in Geochemistry) by Heinrich D. Holland, Princeton Univ. Press, 1984 or Earth : Evolution of a Habitable World by Jonathan I. Lunine, Cynthia J. Lunine, Cambridge Univ. Press, 1998

a very active planet from the geologic point of view. Volcanism was widespread and if we assume that volcanoes five billion years ago emitted the same gases as they do today, the earth's second atmosphere probably consisted of water vapor, carbon dioxide, and nitrogen. These gases were expelled from the earth's interior by a process known as outgassing.

It is also possible that the impact of comets brought significant amounts of water and other volatile gases to the Earth. The vast amounts of water vapor expelled by the volcanic earth resulted in the formation of clouds which, in turn, produced rain. Over a period of thousands of years, the rain accumulated as rivers, lakes and ocean basins. This process was extremely important for the carbon dioxide CO_2. The water reservoirs acted as sinks for that gas and through chemical and later biological processes it became locked up in sedimentary rocks as limestone .

On the other hand nitrogen, which is not very chemically active, continued to accumulate in the atmosphere.

What about the most important gas oxygen we need to live? The first oxidized rocks found in geological strata date back only 1.2 billion years. 600 million years ago oxygen constituted only 1% of the atmosphere (currently 21%). Therefore, Oxygen was only a trace gas in the air when life first appeared on the planet. That was one of the reasons that life first evolved in the oceans. Single-celled bacterium dwelling in the oceans did not need oxygen to live. Oxygen first appeared in the environment when early bacteria developed the ability to split water molecules apart using the energy of sunlight - a key part of photosynthesis. Photosynthesizing organisms produced the oxygen that accumulated over geologic time.

These processes acting sequentially and simultaneously appear to have produced the delicate balance of 78% nitrogen (N_2) and 21% oxygen (O_2) we observe today. By the way, oxygen is the third most abundant element in the universe and makes up nearly half of the mass of the Earth's crust, two thirds of the mass of the human body and nine tenths of the mass of water.

The Earth cannot sustain more than $\sim 20\%\, O_2$ in the atmosphere. Otherwise spontaneous fires would occur that would deplete the oxygen.

The enrichment of oxygen in the atmosphere might be seen in context with the methane content. Microbes who utilize photosynthesis existed on Earth half a billion years or more before oxygen became prevalent, without substantially affecting the composition of the atmosphere. The transition to an atmosphere with noticeable oxygen content occurred about 2.4 billion years ago. According to Catling *et al.* (2001) [55] after photosynthesis separated the oxygen from the hydrogen, the authors argue, the two components followed separate paths. The free oxygen remained in the Earth's crust, while the hydrogen went on to combine with carbon in a process known as "methanogenesis," producing methane. When methane travelled to the upper atmosphere, ultraviolet radiation from the Sun dissolved it into its components. The light hydrogen drifted away into space and was lost forever to the Earth's atmosphere.

Because the hydrogen was lost while the oxygen stayed on Earth, an excess of oxygen gradually accumulated. When the Earth's crust was saturated, the oxygen spilled out and flooded the ancient atmosphere, creating the oxygen rich

environment we know today. This can also solve the faint young Sun problem (see next chapter).

This problem was also recently discussed by Kasting, 2004 [154], where further references can be found. He argues that the climate on Earth prior to 2.5 Gigayears seems to have been even warmer than today, despite the fact, that the luminosity of the Sun was 25-30% less luminous than today (see also next chapter on the faint young Sun problem). Thus a warming of the atmosphere additional to present day greenhouse gases was required and it is argued that this could have been done by methanogens since Ammonia is unstable in low O_2 atmospheres. CH_4 photolyzes only at wavelengths shorter than 145 nm and it is relatively longlived in the absence of O_2, O_3. It is produced by anaerobic bacteria that have evolved early in the Earth's history, the required flux was 500 Tg CH_4/yr. Now there is a positive feedback: this flux should have increased once oxygenic photosynthesis evolved because of increased production and recycling of organic matter. Even if the CH_4 flux would have been the same as at present and the CO_2 at the same level, this would have led to a warming of 30 degrees. However, siderite-coated stream pebbles imply that also the CO_2 concentration was 7 times the present value. A rise in either atmospheric O_2 or oceanic sulfate near the end of the Proterozoic could have caused CH_4 concentrations to decrease a second time and may have triggered the "Snowball Earth" glaciations.

Omori *et al.*, 2004 [238] discuss the role of plate tectonics and the amount of carbon had carried into the mantle via the Archean subduction zone. They found out that plate tectonics can be dated back as early as 3.8 Gigayears.

The role of the changing Sun and its evolution on the Earth's atmosphere was also discussed by Guinan and Ribas, 2002 [121].

The early evolution of the Earth and its atmosphere at the time when the Earth was still growing by planetesimal impacts was studied by Abe and Matsui (1986 [1]). They considered a magma ocean covering the Earth when the accretion time was less than 5×10^7 y. Zahnle *et al.* (1988 [344]) discussed the evolution of an impact generated steam atmosphere. Abe and Matsui (1988 [2]) reported on the evolution of an impact-generated H_2O–CO_2 atmosphere and formation of a hot proto-ocean on Earth.

Chapter 6

Space Weather and Climate

The term Space Weather denotes variations of the Earth's environment on short terms. In analogy to meteorology, where the distinction between weather and climate is made, space climate denotes long term variations of the Earth's climate mainly caused by solar variations.

6.1 The Atmosphere's Response to Solar Irradiation

6.1.1 Introduction

The penetration of solar radiation strongly depends on its wavelength, the larger the wavelength the deeper the penetration (see Fig. 6.1).

The principal effects of solar radiation on the middle and upper atmosphere are summarized in table 6.1. In the second and the third column of that table the variation due to the solar activity cycle is given. From that table it follows that the amount of variation depends on the wavelength of the solar radiation becoming smaller at longer wavelengths. Above 300 nm it is very difficult to detect and can be measured only with satellite radiometric detectors.

In addition to radiation, the Sun also emits the solar wind which consists of particles that interact with the geomagnetic field to form the Earth's magnetosphere. We observe a large input of electrons and protons (causing the aurora) and ionospheric currents are produced causing joule heating. In principle these phenomena are concentrated at high geomagnetic latitudes; heating effects can spread equatorward by convection and conduction.

The typical structure of the Earth's atmosphere was already shortly described. The boundaries of the various layers (Troposphere, Stratosphere, Mesosphere, Ionosphere) are called pauses (e.g. the Tropopause) and are defined by minima or maxima of the temperature profile. At 100 km the density is 10^{-6} of its surface value. The temperature in the thermosphere is strongly dependent on solar activity. The major sources for heating at this layer are:

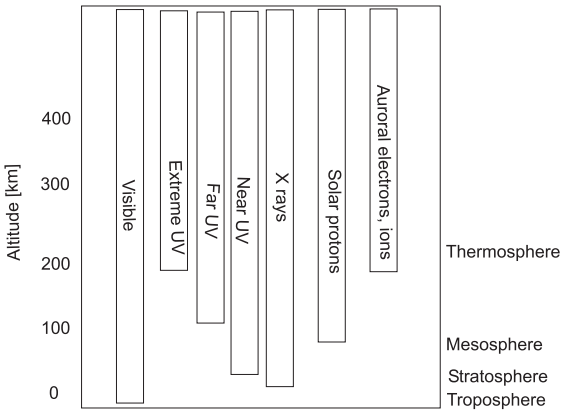


Figure 6.1: Penetration of different solar light waves resp. their induced particles in the atmosphere

Table 6.1: Effects of Solar Radiation at different wavelengths on the Middle and Upper Atmosphere.

Wavelength [nm]	Variab. middle Atm.	Variab. upper Atm.	Effect	Height [km]
1-10, SXR		sporadic	Ion. all	70-100
10-100, XUV	2ppm	2 x	Ion. N_2, O, O_2	
100-120, EUV	6 ppm	30%	Ion. NO	80-100
120-200, VUV	150 ppm	10%	Diss. O_2	40-130
200-240, UV	0.12%	5%	Diss. O_2, O_3	20-40
240-300, UV	1.0%	<1%	Diss. O_3	20-40

- solar ionizing photons,

- magnetospheric processes.

The principal part of the ionosphere is produced by XUV which is strongly absorbed there. Ionization and recombination occurs and this contributes to the heating of the thermosphere. By comparing with table 6.1 we see that the energy involved is small; dissociation of O_2 is strong, above 120 km oxygen occurs as atoms. Through vertical mixing the ratio $O_2 : N_2$ is constant near 0.1 throughout the lower thermosphere. Oxygen atoms are produced down to 30 km and most of them combine with O_2 to form ozone. This attains a peak ratio of 10^{-5} near 30 km. Photons around 300 nm can reach the surface. They produce electronically excited oxygen and surface ozone which drives a large fraction of urban pollution chemistry.

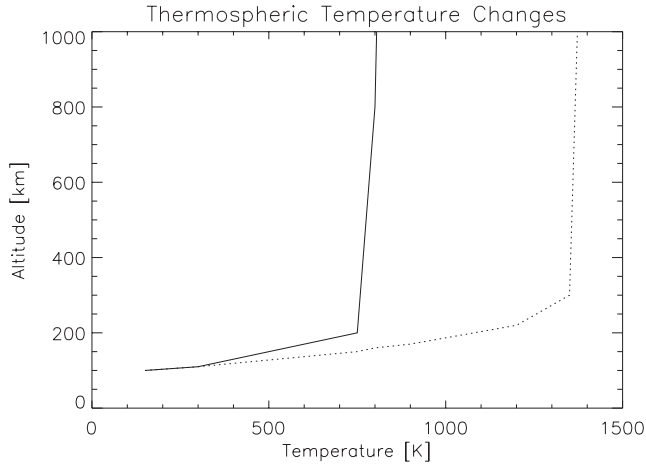


Figure 6.2: Thermospheric temperature changes, a) low solar activity, $F_{10.7} = 80, A_p = 0$, b) high solar activity, $F_{10.7} = 200, A_p = 80$ (dashed line). F10.7 denotes the 2800 MHz solar flux which is a measure for solar activity, the Ap index is a measure of the general level of geomagnetic activity over the globe. Up to the top of the troposphere the two curves are nearly identical. At a height of about 100 km the difference is already of the order of several 10 K.

6.1.2 UV Radiation

Solar radiation shortward 320 nm represents only 2% of the total solar irradiance; 0.01% of the incident flux is absorbed in the thermosphere at about 80 km and 0.2% in the stratosphere above 50 km. This radiation is extremely important since the thermal structure and photochemical processes above the troposphere are controlled by it. The stratosphere is controlled by absorption and dissociation of O_2 in the 175 to 240 nm range. The 205 to 295 nm range is predominantly absorbed by ozone O_3. If there is a stratosphere- troposphere coupling, this could affect also the climate. The short term variation of UV radiation is ascribed to the evolution and rotation of plage regions on the solar disk. The XUV induced thermospheric temperature changes is shown for low and high solar activity in Fig. 6.2. Solar activity is measured in terms of the 10.7 cm radio flux and of the plage area A_p.

6.1.3 Energetic Particles

There are three main contributions:

- electrons: they reach the high latitude thermosphere after interaction with the geomagnetic field and acceleration;
- high energy solar protons: their flux is enhanced during periods of large flares;

Table 6.2: Exospheric temperature at solar maximum and minimum

	Temperature of exosphere
solar minimum	700 K
solar maximum	1200-1500 K

- galactic cosmic rays: they originate from outside the heliosphere but their input on Earth is partly controlled by solar activity.

During large flares, intense fluxes of energetic protons (10...10^4 MeV) penetrate the Earth's polar cap regions. They produce ionization between 100 and 20 km. Such an event can last for a few hours to a few days. Large numbers of NO_x molecules are produced leading to a subsequent ozone depletion.

Relativistic electron precipitation are possible sources for ionization and odd nitrogen production at altitudes above 80 km, thus well above the ozone layer.

6.1.4 Thermosphere and Exosphere

The thermosphere starts at a height of about 90 km and ends at about 250 km at the so called exobase. In the thermosphere the temperature gradient is positive, in the exosphere collisions become negligible, particles execute ballistic orbits. The most important heat source is solar XUV radiation creating the ionosphere. The resulting heating is conducted down to the mesopause where it can be radiated. The exosphere is approximately isothermal because it lies above the level where most of the energy is deposited. Also the thermal time constant is very short because of the low density there.

The variations can be divided into diurnal variations and longer term variations.

- The diurnal variations show a day/night ratio of 1.28 over the equator with the peak occurring about 2 p.m. During the night, heat is conducted down from the top of the thermosphere to its base, the mesopause, where it is radiated.

  The effect of the sun can be expressed and measured by the 10.7 cm radio flux which is given in the units $10^{-22}\,\mathrm{Wm^{-2}Hz^{-1}}$. If this quantity is multiplied by a factor of 1.8 deg per unit of flux, we obtain the temperature.

- The other variation comes from the solar activity cycle (see Table 6.2).

The thermospheric temperature changes are illustrated in Fig. 6.2. Similarly the peak electron density varies by a factor of 2 which is very important for short-wave radio communication. The temperature changes are obtained from the 10.7 index with a multiplier of 3.6. This value is twice than that for the 27 day variation which is due to solar rotation. The difference may result from the fact that in the case of activity cycle variations not only the variation of XUV but also a

contribution from auroral heating- which is triggered by the solar wind- must be taken into account.

Finally we have to stress that both UV and particle precipitation have chemical effects and the most important is the production of N, NO and NO_2 (which is collectively called NO_x). The following reactions define the NO_x production in the thermosphere:

$$O^+ + N_2 \rightarrow NO^+ + N \quad (6.1)$$
$$N_2^+ + O \rightarrow NO^+ + N \quad (6.2)$$
$$N_2^+ + e^- \rightarrow N + N \quad (6.3)$$
$$NO^+ + e^- \rightarrow N + O \quad (6.4)$$

6.1.5 Mesosphere and Stratosphere

This region extends from the tropopause to the mesopause, at approximately 90 km. It is in local radiative equilibrium except from heat flowing in from the thermosphere. The radiative heating is by absorption of planetary radiation mainly by:

- CO_2 at 15 μm,
- O_3 at 6.3 μm
- absorption of solar UV by O_3.

The CO_2 band is the principal radiator. The O_3/CO_2 ratio decreases upwards; thus the heating to cooling also decreases and the temperature gradient is negative. In the stratosphere ozone begins to be more and more attenuated, a temperature maximum at 50 km occurs, the stratopause.

The solar UV flux is not variable at large scales thus the temperature changes to be expected from a variation of that flux should also be small. Most effects therefore come from ionization and a changing chemistry. The changes of the UV flux from the sun have only a modest effect on ozone amounts because both production (by a photolysis of O_3) and destruction are affected in the same way. Another effect is the penetration of solar protons or relativistic electrons into the middle atmosphere. By that penetration considerable amounts of NO_x are produced; these enhancements of NO_x increase the destruction of ozone at high altitudes. This could explain the inverse correlation of ozone amounts with solar activity found by Ruderman and Chamberlain (1975 [264]). Chakrabarty, 1982 [61] studied how Ozone is affected by solar proton events. During such events NO is produced that destroys ozone.

To test these predictions it is important to have data at the time scale of the solar cycle; however we must also take into account the instrumental drifts as well as the typical lifetime of the instruments which normally are below 5 yr. In the stratosphere, the ozone response is caused primarily by changes in production from O_2 and has a maximum value of 0.5 % for a 1 % change in the UV at 205 nm. The study of the response of the temperature has been made by Hood (1986

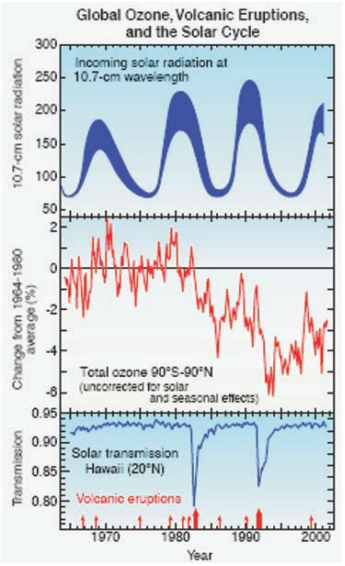


Figure 6.3: Relation between global ozone, solar activity (measured by 10.7 cm radiation) and volcanic eruptions (El Chichon, 1982, Mt. Pinatubo, 1991). Large volcanic eruptions decrease solar transmission and the particles enhance ozone depletion. Adapted from D.W.Fahey *http* : *//www.epa.gov/ozone/science/unep*5*ciQandA.pdf*

[136], 1987 [137]) and Keating *et al.* (1987 [157]) between 30 and 0.2 mbar (24 to 60 km) and later by Clancy and Rusch (1989 [68]) up to 90 km. They establish the already mentioned 0.5 % response. The very small temperature response lags the UV by 4 to 14 days.

The response of the stratospheric ozone concentration to solar activity was examined in different studies. A study by Angell and Korshover (1976 [11]) established a correlation of the ozone column with solar activity with a peak to peak variation up to 10 % at 70^0 latitude and only 4 % at 47^0. In the upper stratosphere there is an increasing trend of ozone with increasing solar activity. Enhanced photolysis of NO however reduces the ozone destruction during solar activity maximum. An ozone variation of 3% seems to correspond to a 20% solar UV variability at 180 nm (Keating, 1981 [156]). In general the stratospheric ozone content varies with the solar activity cycle whereas the tropospheric ozone does not. The stratospheric ozone decreases 2.72% - 3.79%, and total ozone 2,71 % - 4.36% when solar activity decreases; when solar activity increases the stratospheric ozone increases 2.41% - 3.06% and total ozone increases 2.1% - 5.56% (see Asiati *et al.*, 2004 [15]).

There seems to be no correlation of polar stratospheric temperatures and solar activity (Labitzke (1987 [179]), Labitzke, Van Loon (1988 [180]), Kerr (1988 [159]). There exists a stratospheric biennial oscillation which is more or less periodic and reversal of winds in the lower equatorial stratosphere with an average period of 27 months.

6.1.6 Troposphere

As we have seen above, only wavelengths > 300 nm penetrate to the troposphere and surface. We have already stressed that this part of the solar spectrum is only slightly variable with a peak to peak variation of about 1 part in 1400. Thus the troposphere which contains 90 % of the total mass of the Earth's atmosphere is subject to a nearly constant driving solar energy.

However, there have been innumerable attempts to find correlations between solar activity and various meteorological phenomena and other variables. If the troposphere is to be significantly influenced by the tiny changes of solar irradiation, there should exist a very strong mechanism of amplification (trigger mechanism). Such mechanisms were discussed:

- magnetospheric effects by electric field - including also effects of thunderstorms (Mc Cormac and Seliga, 1979 [215]).
- Hines (1974 [134]) suggested a change of the transmissivity of the stratosphere to upwardly propagating atmospheric waves (Callis *et al.* 1985 [52] showed from models that this is possibly not the case).
- The effect found by Labitzke (1987 [179]): temperatures in the polar winter are jointly influenced by the solar cycle and the quasi biennial oscillation and the effect on the troposphere is discussed in Van Loon and Labitzke (1988 [180]).
- Eddy (1976 [87], 1988 [88]) discussed the absence of sunspot activity during the 17th century which is known as the Maunder minimum and an earlier event, called the Spörer minimum. Both periods seem to coincide with periods of reduced global temperatures the more recent is called the Little Ice Age. Eddy (1988 [88]) showed that the required solar input reduction would have to be much greater than the tiny amplitudes detected on the time scale of a solar cycle. Maybe also amplifying factors have to be considered.

6.2 The Faint Young Sun

6.2.1 Evolution of the Solar Luminosity

According to theories of stellar evolution, the solar constant[1] is not a constant but has been increasing continuously throughout the main sequence lifetime of the Sun. The increase in luminosity can be explained by the process of energy generation inside the Sun, the nuclear fusion of hydrogen into helium; by this energy generation the mean atomic weight and density of the Sun is increased. Since the gas pressure is given by

$$P \approx kT/\mu \tag{6.5}$$

[1]The amount of energy from the Sun received per unit area at the Earth

an increase of the molecular weight μ implies a higher temperature T in order to sustain the gravity and to keep hydrostatic equilibrium. An increase in T means an increase in the energy production and thus luminosity L.

A very rough formula for the luminosity change of the Sun during its main sequence evolution was given by Gough (1981) [118]:

$$L(t) = [1 + 0.4(1 - t/t_0)]^{-1} L_0 \tag{6.6}$$

In this formula L_0 is the present solar luminosity and t_0 the present age of the Sun (4.6 Gyr). Other explanations of a possible different solar luminosity at the early evolution of the Sun are:

- Revisions in the standard solar model in order to solve the neutrino problem.
- Strong mass loss during the early phase (Willson *et al.* 1987 [335]).

Sagan and Mullen (1972) first pointed out the implications of this change of solar luminosity for the Earth's climate[2]. Using a very simple model of the greenhouse effect they showed that lower solar luminosity would have resulted in T_s below the freezing point of water for roughly the first 2 Gyr of the Earth's evolution. However this cannot be correct. Already Sagan and Mullen pointed out the presence of pillow lavas, mud cracks and ripple marks in 3.2 Gyr old rocks suggesting strongly the presence of liquid water on the Earth's surface at that time. We also know that sedimentary rocks have been deposited about 3.8 Gyrs ago and these must have formed in liquid water.

We should also stress here, that the faint young Sun was more active and variable than today, especially in the short wavelengths (X, UV).

6.2.2 Pre Main Sequence Sun

All calculations show that during the early life of the Sun, the UV flux was much higher than today. The Sun had a behavior similar to a T Tauri star. Zahnle and Walker (1982 [344]) calculated that the flux decreases as

$$\sim t^{-s} \qquad 0.5 < s < 1 \tag{6.7}$$

The exponent in this formula depends on the wavelength considered. Similar results were obtained by Canuto *et al.* (1982).

6.2.3 Albedo Variations

Let us consider the Earth to radiate like a blackbody, S is the solar constant (at present 1360 W/m^2), σ the Stefan-Boltzmann constant , A the planetary Albedo (~ 0.3), T_e the effective radiating temperature can be obtained by:

$$T_e = [S(1 - A)/4\sigma]^{1/4} \tag{6.8}$$

[2]see also e.g. The Role of the Sun in Climate Change by Douglas V. Hoyt, Kenneth H. Schatten, Kenneth H. Schatten, Oxford Univ. Press, 1997

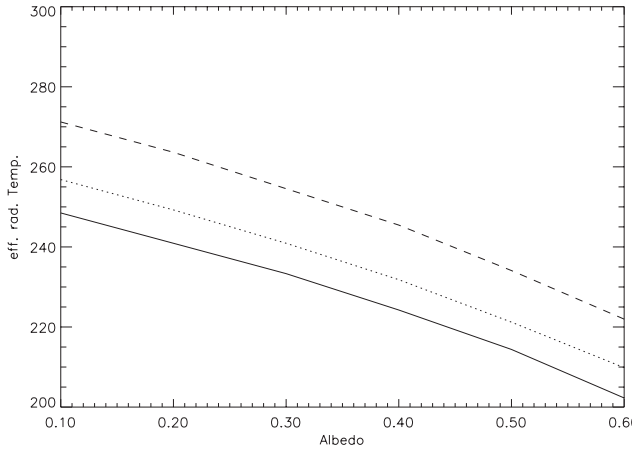


Figure 6.4: Effective radiating temperature of the Earth as a function of planetary Albedo A for three different values of the solar constant, a) 982, b) 1088 (dotted), c) present value 1360 (dashed).

The relevant albedo to use here is the *Bond Albedo*, which is the percentage of the total incident solar radiation over radiation reflected back into space. The present effective radiating temperature of the Earth is $\sim 255\,\mathrm{K}$. If we combine 6.6 and 6.8 then the increase of T_e was about 20 deg over geologic time if the albedo of the Earth is assumed to remain constant. We must also take into account the Earth's mean surface temperature T_s and

$$T_\mathrm{s} > T_\mathrm{e} \tag{6.9}$$

Because of the greenhouse effect the difference between T_s and T_e is about 33 K. The greenhouse effect is caused by the difference in opacity in the visible and infrared regions of the electromagnetic spectrum. The Earth's atmosphere is relatively transparent to incoming solar radiation, but absorbs a large fraction of the outgoing IR. Most of the absorption is caused by the vibration-rotation bands of H_2O and CO_2 and to the pure rotation band of H_2O.

In Figure 6.4 the effective radiating temperature of the earth T_e as a function of planetary albedo is given for three different values of the solar constant, a) at present b) reduced by 20 % and c) reduced by 30 %.

We clearly see that T_e strongly depends on the solar constant and on the Albedo. A larger value of the albedo leads to a lower value of the effective radiating temperature, the Earth becomes cooler. The albedo can increase because of:

- increased glaciation of the Earth,
- increased fraction of clouds.

Some typical values for A are given in Table 6.3.

Table 6.3: Typical values for the albedo.

	Albedo
Tropical forest	0.13
Woodland	0.14
Grassland	0.20
Stony desert	0.24
Sandy desert	0.37
Sea ice	0.25-0.60
Snowy ice	0.80
Cool water	<0.08
Warm water	< 0.10

One possible explanation might be that the Earth's albedo was significantly lower in the past or that the greenhouse effect of its atmosphere was larger. However, as Sagan and Mullen pointed out a large change in the Earth's albedo was unlikely; any decrease in cloudiness that might result from lower surface temperatures would likely be compensated by an increase in snow and ice cover. However if the Earth's surface was mostly water covered this argument does not work.

From climate research we know that there was no glaciation on Earth prior to about 2.7 Gyr ago (e.g. oxygen isotopes imply warm surface temperatures throughout the Precambrian (Kasting and Toon, 1989 [155])).

6.2.4 The CO_2 Geochemical Cycle

Let us start with some numbers: the total surface reservoir of carbon is about 10^{23} g. This is enough to produce a CO_2 partial pressure of about 60 bar were all of it present as gaseous (Holland, 1978). Most of the carbon is contained in carbonate rocks on the continents. A much smaller amount is present in the oceans as carbonate and bicarbonate ions (4×10^{19} g). Presently about 7×10^{17} g are present in the atmosphere (this number is growing). There is an equilibrium between the ocean and the atmosphere at timescales of about 1000 y:

CO_2 is removed from the atmosphere/ocean by weathering of silicate rocks on the continents and 20 % of atmospheric CO_2 is removed by photosynthesis followed by burial of organic carbon. If one represents silicate rocks by $CaSiO_3$ (wollastonite) then the CO_2 loss process can be described by the following three reactions:

$$CaSiO_3 + 2CO_2 + H_2O \rightarrow Ca^{++} + 2HCO_3 + SiO_2 \quad (6.10)$$

$$Ca^{++} + 2HCO_3 \rightarrow CaCO_3 + CO_2 + H_2O \quad (6.11)$$

$$(6.12)$$

Thus:

$$CaSiO_3 + CO_2 \rightarrow CaCO_3 + SiO_2 \quad (6.13)$$

When old sea floor is subducted and carbonate sediments are subjected to higher temperatures and pressures CO_2 is returned to the atmosphere/ocean. Then reaction 6.13 goes in the opposite direction, calcium silicate is reformed and gaseous CO_2 is released. Much of this CO_2 escapes through volcanoes. That process is termed carbonate metamorphism and on the young Earth the rate of carbonate metamorphism could have been augmented by faster rates of tectonic cycling and by impact processing of carbonate rich sediments.

It is important to note that the rates of the weathering reactions are strongly dependent on temperature. The reaction rates increase with temperature and weathering requires liquid water. The temperature dependence of the silicate weathering process rate leads to a negative feedback between atmospheric CO_2 and surface temperature: if the surface temperature were to decrease (because of a faint young Sun), the weathering rate would also decrease and carbon dioxide would begin to accumulate in the atmosphere. This increase of CO_2 causes an increase in the greenhouse effect and thus the temperature increases. The reverse would happen if the climate became warmer: the weathering rate would increase, pCO_2 would decrease and the greenhouse effect would become smaller (Walker *et al.* 1981 [328]). This mechanism can explain why the temperature on Earth was high enough for liquid water even when the solar luminosity was smaller.

The modern rate of CO_2 release from volcanoes would create a 1-bar CO_2 atmosphere in only 20 Myr if carbonates were not forming. This shows that the response time of the system is quite fast in geologic terms.

6.2.5 Effects of the Biota

Presently the CO_2 geochemical cycle is modulated by the biota. Calcium carbonate formation can be largely attributed to the secretion of shells by plankton and other marine organisms. Land plants enhance silicate weathering rates by pumping up the carbon dioxide partial pressure in soils by a factor of 10 to 40 over the atmospheric value; photosynthesis on land and in the oceans creates organic carbon which is then buried in sediments. Thus the atmospheric CO_2 level is reduced. Therefore, the Earth today is probably cooler than it would be in the absence of life. Lovelock (1988 [201]) created therefore the Gaia Hypothesis, which means that the Earth's climate is controlled by biota and would have become unstable where it not for the homeostatic modulation of climate by organisms. Let us assume biological control of the Earth's climate in more detail. According to Berner *et al.* [35], the dependence of the silicate weathering rate f_w on surface temperature T can be written as:

$$f_w = 1 + 0.087(T - T_0) + 0.0019(T - T_0)^2 \tag{6.14}$$

Here, T_0 is the present mean surface temperature (288 K) and $f_w = 1$. The CO_2 greenhouse effect is parameterized in the BLAG model as:

$$T - T_0 = 2.88 \ln(P/P_0) \tag{6.15}$$

where P indicates atmospheric pCO_2 and P_0 the present CO_2 partial pressure. From laboratory studies we know that the weathering rate of silicate minerals

varies approximately as

$$\mathrm{pCO}_2^{0.3} \tag{6.16}$$

for CO_2 partial pressures of 2 to 20 bar and temperatures of 100 to 200^0 C. These data were derived by Lagache (1976 [183]) and Walker *et al.* (1981 [328]). Let us assume that we can apply this relation to the Earth's surface conditions. To study the maximum effect let us further assume that removing land plants from the system would reduce surface soil pCO_2 by a factor of 40. Then the equation for the silicate weathering process can be written as:

$$\begin{aligned} f_{\mathrm{w}} &= [1 + 0.087(T - T_0) + 0.0019(T - T_0)^2] \\ &= [P_{\mathrm{S}}/40P_0)]^{0.3} \end{aligned} \tag{6.17}$$

Here P_S is the partial CO_2 pressure in the soil and today we have $P_S = 40P_0$ and obtain f_w=1. On a vegetation free Earth $P_S \sim P_0$, and f_w would be reduced by a factor of $40^{-0.3} \sim 1/3$. The carbon cycle is only balanced when $f_w = 1$. Therefore, without vegetation, the atmospheric pCO_2 and surface temperature would have to increase to bring back the silicate weathering rate to its present value. We substitute equation 6.15 into equation 6.17 and solve for P/P_0 and obtain:

$$P/P_0 = 9 \qquad T - T_0 \sim 6.3\,\mathrm{K} \tag{6.18}$$

This shows that under the assumption that land plants pump up soil CO_2 by a factor of 40, the effect of eliminating them would be to increase the Earth's temperature by only 6 deg. The net cooling effect of the biota should be somewhat larger because of the influence of the organic carbon cycle; today 20% of the carbon is organic carbon rather than carbonate. One can estimate that if life suddenly were eliminated in total the temperature would increase by 8 deg. Thus even a lifeless Earth would apparently be no warmer than the real Earth was during the Cretaceous, when the dinosaurs flourished.

The studies of Schwartzmann and Volk (1981) [275] showed that biota may accelerate chemical weathering by stabilizing soil (silicate minerals stay in contact with carbonated water), generating organic acids. This could lead to enhanced weathering rates of up to 1000 instead of 3. Therefore, the CO_2 partial pressure on a lifeless Earth might be as high as a few tenths of a bar and the surface temperature may be up to 60 K warmer!

6.2.6 T Tauri and Post T Tauri Phase

T Tauri stars are a group of stars which are solar like and often associated with molecular clouds. They are very early stars that means that they have not yet reached the main sequence and they are still contracting (see the introduction about stellar evolution). Their masses and temperatures are quite similar to the Sun but they are brighter. Their rotation rate is in the range of a few days (for the Sun it is about 1 month). They are active variable stars. The first ones were found about 1945 by their optical variability and chromospheric lines. Later on some evidence for large starspots on their surfaces were found. The X ray emission

which is about 1 000 times that of the present Sun and radio flux is not constant. Some of them also show molecular outflow and strong stellar winds. By their IR and sub mm excess radiation it was found that about half of them are surrounded by circumstellar disks.

Contrary to normal main sequence stars like the Sun their energy is not produced by nuclear fusion near the core but by a slow gravitational contraction. T Tauri stars belong to the group of so called YSO (young stellar objects) of type II. Type I YSO are very young protostellar objects at the age of just a few 100 000 years. An example is HR 4796. At a wavelength of 12 μm the object appears as a point source, at a wavelength of 21 μm it is much larger and diffuse indicating a circumstellar disk of dust. Such objects can be observed preferentially in the IR.

At an age of about 40 Million years our Sun became a zero age main sequence star. That means that it reached the main sequence and nuclear fusion of H to He started. At that time the Sun had about 70% of the total flux that is emitted presently. But in the UV and X-rays the flux was higher by a factor of about 100 than now. This of course has important consequences for the formation of the planets, their atmospheres etc. In its T Tauri and post T Tauri evolution the Sun's short wavelength emission was considerably higher than it is now. At that time the terrestrial planets were formed already, the protoplanetary disk evaporated, comets ejected out into the Oort cloud and the big bombardment period from the remaining rocky planetesimals and comets began; this caused probably several evaporations of the Earth's oceans.

How can we find indications for this T Tauri and Post T Tauri Phase of the Sun? Measurements of the ^{15}N to ^{14}N ratio in the atmosphere of the satellite Titan (which is Saturn's largest satellite) have shown that the bulk N is enhanced in the heavier ^{15}N isotope by about 4.5 times relative to the Earth's value. A ^{15}N/^{14}N anomaly on Mars of about 1.6 times the terrestrial value has also been found. These measurements can only be explained by the above mentioned T Tauri and post T Tauri phase of the early Sun (Lammer *et al.* 2000 [186]). Atmospheric sputtering and pick-up caused by a high solar wind particle outflow during a Post T-Tauri phase could be responsible for the observed nitrogen anomaly.

6.3 Solar Variability

Solar variability can be divided into three components according to their influence on the structure and composition of the atmosphere:

- variation of the solar constant
- XUV and UV variation
- energetic particle variation

So far we have only discussed the long term solar variability- summarized as the faint young Sun problem and the influence of the changing parameters of the Earth's orbit on climate (Berger, 1980 [34]).

Table 6.4: Satellite measurements of the solar constant

NIMBUS-7	16 Nov 78-13 Dec 93
SMM (ACRIM I)	16 Feb 80-01 Jun 89
ERBS	25 Oct 84-
NOAA-9	23 Jan 85-20 Dec 89
NOAA-10	22 Oct 86-01 Apr 87
UARS (ACRIM II)	5 Oct 91-
SOHO/VIRGO	18 Jan 96-

The solar input can be represented by the already mentioned solar constant S. In this context we give two interesting facts: First suppose that the Sun has no nuclear sources then the present luminosity could be maintained by a gravitational contraction of only 10^{-4} arcsec/year. Such a change would have been imperceptible over historical times. The other fact is that, as we have seen convection is a stochastic process and it is transport energy to the surface. The efficiency of convection is given by

$$l/H \tag{6.19}$$

where l is the mixing length and H the pressure scale height. Dearborn and Newman (1978) [73] show that a variation of l/H by 0.02 causes a variation of $\Delta S/S$ by 1%. Such a variation is assumed to change the global temperature on Earth by 2 K.

We now address to the question whether there exists also a variability of the solar input on shorter timescales.

6.3.1 Total Solar Irradiance Measurements

The total solar irradiance describes the radiant energy emitted by the sun over all wavelengths that falls vertically each second on 1 square meter outside the earth's atmosphere. This is the definition of the solar constant. Because of the influences of the Earth's atmosphere this constant is extremely difficult to measure on the surface and the most reliable measurements can only be done from space. In Table 6.4 the satellite measurements and the respective time spans of the measurements are summarized.

The VIRGO Experiment on the ESA/NASA SOHO Mission has two types of radiometers to measure total solar irradiance (TSI): DIARAD and PMO6V. A description of the instrument can be found in Fröhlich *et al.* (1995) [105]. Let us shortly describe the DIARAD measurement facility which is a part of SOHO/VIRGO: DIARAD is a Differential Absolute Radiometer. It is composed of two cylindrical cavities coated inside with diffuse black and mounted next to each other on the same heat sink. The flat bottom of the cavities are in fact heat flux transducers on which heating elements have been mounted. Both cavities see the same thermal environment through accurately known circular apertures.

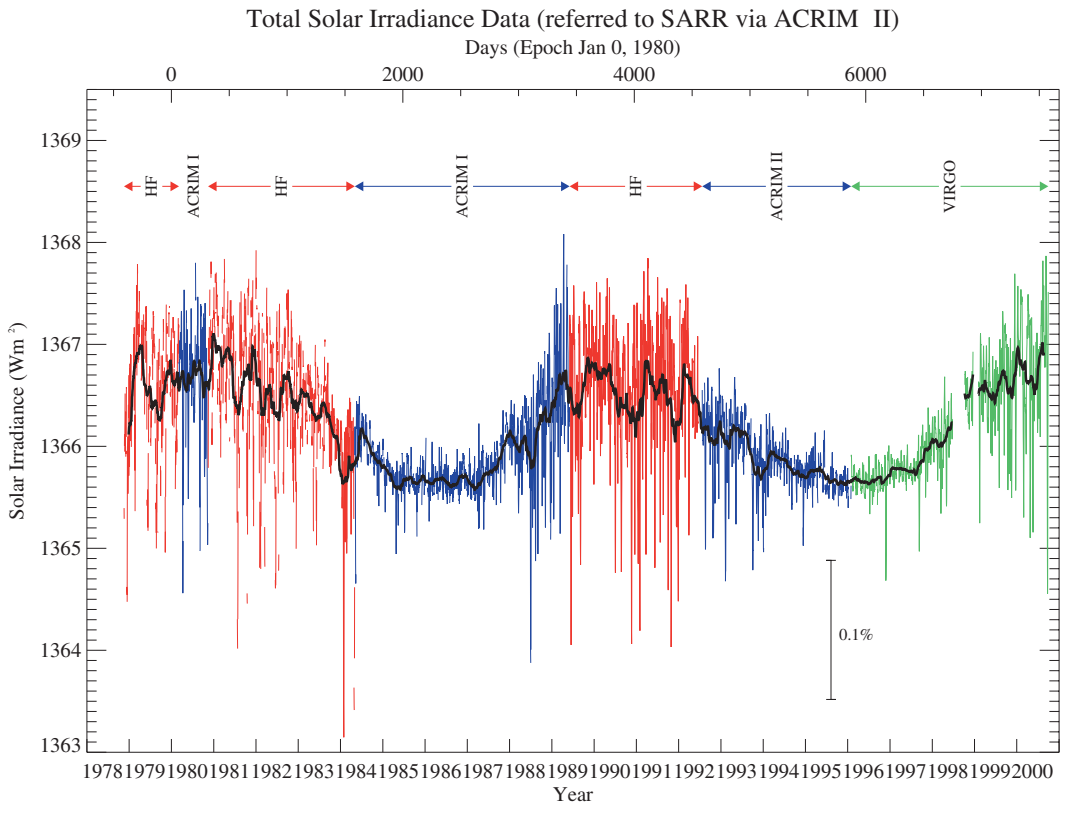


Figure 6.5: Solar irradiance measurements from satellites

A comparison of the power generated inside the cavities is done. For instance a constant electrical power is generated in one of the channels and the difference between the two heatflux sensors is automatically brought back to zero by an ad hoc accurate servosystem that provides electrical power to the other channel called "active channel". This one is regularly irradiated by the Sun or closed. The difference of the electrical power fed to the active channel when its shutter is open (exposed to the Sun) and when it is closed is proportional to the incident solar irradiance. From time to time, the roles of the left channel and the right channel are reversed for half an hour with the purpose of monitoring the aging of the continuously exposed left channel. The sampling rate of the PMO6 instrument is 1 solar total irradiance / 2 minutes, for DIARAD 1 solar total irradiance / 3 minutes.

The ACRIM contains four cylindrical bays. Three of the bays house independent heat detectors, called pyrheliometers, which are independently shuttered, self calibrating, automatically controlled, and which are uniformly sensitive from the extreme UV to the far infrared. Each pyrheliometer consists of two cavities, and temperature differences between the two are used to determine the total solar flux. One cavity is maintained at a constant reference temperature, while the other is heated 0.5 K higher than the reference cavity and is exposed to the Sun periodically. When the shutter covering the second cavity is open, sunlight enters, creating an even greater difference in cavity temperatures. The power supplied to the second cavity by the ACRIM electronics decreases automatically to maintain the 0.5 K temperature difference between the two cavities. This decrease in the amount of electricity is proportional to the solar irradiance entering the cavity. Exposing the sensors to the space environment and the Solar UV radiation causes some small changes on the surface of the cavities which may affect the measurements. The ACRIM instrument monitors this type of problem by carrying three similar sensors, two of which are normally covered. At times these are opened for comparison purposes. Further details can be found in Willson (1981 [336], 1984 [337]).

Measuring the solar constant one finds:

- Part of the energy is blocked by dark sunspots and subsequently released in faculae. The screening effect by sunspots is overcompensated by the energy storage and release. This is demonstrated in Fig. 6.6.
- There are variations of the solar constant with the solar cycle.

First measurements with the ACRIM 1 (Active Cavity Radiometer Irradiance Monitor) experiment on board the Solar Maximum Mission and the ERB experiment on the Nimbus-7 satellite showed a positive correlation between the solar cycle activity, measured by the sunspot index, and the total solar irradiance. The peak to peak variation of about 1 W/m^2 (out of about 1367) between solar maximum and minimum was reported by Fröhlich (1987) [104], Willson and Hudson (1988) [338] and Foukal and Lean (1988) [100]. Somewhat larger fluctuation up to 0.2% occur over timescales of days and weeks.

Given that the total variation between the peaks of solar cycles 21 and 22 was about 0.1%, how much is the effect to be expected for a change of the cor-

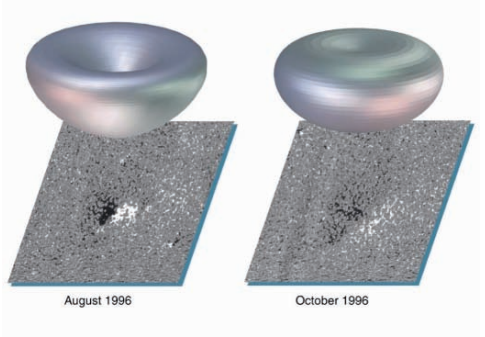


Figure 6.6: Three-dimensional rendering of the angular distribution of the excess irradiance emitted at 500 nm by the active region studied at two stages of its development, together with the magnetogram. A more uniform brightening of the facular region at the later stage is apparent (after Vicente Domingo).

responding global temperature on Earth? It is expected that this change of the solar irradiation produces a corresponding variation of about 0.2^0 C in globally averaged equilibrium surface temperature (Hansen and Lacis, 1990 [125]). But there is some considerable delay in the response. Because of the thermal inertia of the oceans, the time needed to approach equilibrium is much longer than 11 years (e.g., Reid, 1991 [254]), so that the actual temperature response to the observed variation during a solar cycle is likely to be considerably smaller, and probably insignificant from a climatic point of view.

The climate sensitivity of solar variability was studied by Scafetta and West, 2004 [271]. They estimated a solar net forcing between 10-30% of the global temperature trend between 1980 and 2002.

The complex interaction of solar variation on the atmosphere was modelled (Knipp *et al.* 2004 [169]) by using three input data:

- The solar extreme UV (XUV) power
- Joule power
- Particle kinetic power

From such a model both influences a) radiation, b) particles from the sun to the Earth can be taken into account. During 1975 and 2003 the contributions were found to be: particles 36 GW, Joule 95 GW and 464 GW for the rest. Solar wind-driven geomagnetic power provided 22% of the total global upper atmospheric energy. An interesting trend is that with increasing activity (here by the term activity short time scaled events are meant) the Joule power becomes more and more important. In the top 15 power events, geomagnetic activity contributed to 2/3 of the total power budget.

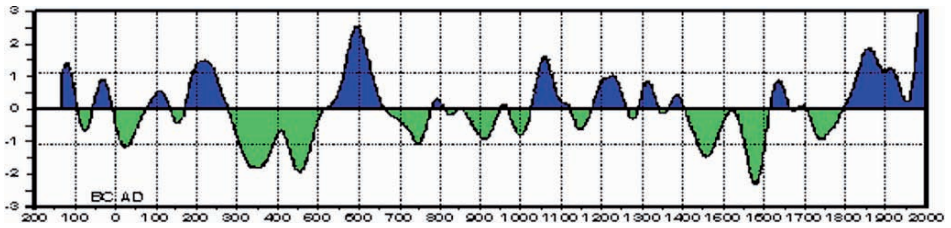


Figure 6.7: Reconstructed precipitation in northern New Mexico. Courtesy: Henri D Grissino-Mayer

6.3.2 Long Term Solar Variations

In order to study a long term variation of the solar output, there is no direct observational support. It is therefore necessary to use proxy data or solar activity indicators. Sunspot index measurements exist over a time span of roughly 350 years and they suggest the presence of a 76-80 yr cycle, the Gleissberg cycle, modulating the 11 yr cycle (Sonett, 1982 [294], Berry, 1987 [36]). Foukal and Lean (1990 [101]) gave an empirical model of total solar irradiance variation between 1874 and 1988.

The presence of the solar cycle has been claimed in various sets of proxies:

- auroral activity,
- isotopic composition of ice cores (Johnsen *et al.* 1970 [149]),
- tree growth, dendroclimatic investigations (Svenonius and Olausson, 1979 [304]). For annual rings to form, trees must "shut down" growth at some point to form a distinct ring boundary. This occurs in the dormant season, usually in the fall and winter. In the tropics, the seasons are not as distinct, so that trees can grow year-round. One fundamental principle of dendrochronology is "the present is the key to the past," originally stated by James Hutton in 1785. However, dendrochronology adds a new "twist" to this principle: "the past is the key to the future." In other words, by knowing environmental conditions that operated in the past (by analyzing such conditions in tree rings), we can better predict and/or manage such environmental conditions in the future. Hence, by knowing what the climate-tree growth relationship is in the 20th century, we can reconstruct climate from tree rings well before weather records were ever kept! Let us give one example from Grissino-Mayer: Fig 6.7 shows a long-term precipitation reconstruction for northern New Mexico based on tree rings. How this reconstruction was made? The reconstruction was developed by calibrating the widths of tree rings from the 1900s with rainfall records from the 1900s. Because we assume that conditions must have been similar in the past, we can then use the widths of tree rings as a proxy (or substitute) for actual rainfall amounts prior to the historical record.

Individual tree-growth series can be "decomposed" into an aggregate of environmental factors:

$$R_t = A_t + C_t + \Delta D1_t + \Delta D2_t + E_t \tag{6.20}$$

R_t is the tree ring growth as a function of t. A_t age related growth trend due to normal physiological aging processes, the climate (C) that occurred during that year the occurrence of disturbance factors within the forest stand (for example, a blow down of trees), indicated by $D1$, the occurrence of disturbance factors from outside the forest stand (for example, an insect outbreak that defoliates the trees, causing growth reduction), indicated by $D2$, and random (error) processes E not accounted for.

A study of tree rings and application to reconstruct climate was given by Cook *et al.* (1997). Sampling 300-to-500-year-old Siberian pine trees in the Tarvagatay Mountains of western central Mongolia, D'Arrigo *et al.*(1993 [71]) analyzed annual growth rings, which generally grow wider during warm periods and narrower in colder times in trees at the timber line. They developed a tree-ring record reflecting annual temperatures in the region dating back to 1550. The Mongolian tree rings show temperature changes that are strikingly similar to records from tree rings in North America, Europe and western Russia. The general trends reflected in the tree-ring record include cooler conditions in the early 1700s, followed by warming that started mid-century. An abrupt cooling occurred in the late 1700s and continued for much of the 1800s. The coldest period was between 1830s and 1870s, after which a steadily increasing warming trend began. An example of this analysis is given in Fig 6.8.

- Solar radius variations (Gilliland, 1981 [111]) ,
- sedimentary rocks (Sonett and Williams, 1985 [295])
- sea surface temperatures (Gerard, 1990 [107]). The mechanism how this could be related to solar irradiance variations works as follows:

  1. absorption of solar energy by the tropical oceans in a deep surface layer,
  2. transport of that energy by ocean currents,
  3. transfer of that energy by evaporation into atmospheric moisture and pressure systems leading to more precipitation (Perry, 1994 [244]).

Lewis *et al.* (1990 [195]) showed that solar radiation in visible frequencies, usually assumed to be absorbed at the sea surface, penetrates to a significant depth below the upper mixed layer of the ocean that interacts directly with the atmosphere. In clear water, the blue wavelengths, where the greatest amount of energy is available, penetrate the deepest, to nearly 100 m. Energy injected into the ocean at this depth can be stored for a substantial period of time.

As it has been stated above the transparency of the tropical oceans is dependent upon the amount of biogenic material, phytoplankton pigments, and degradation

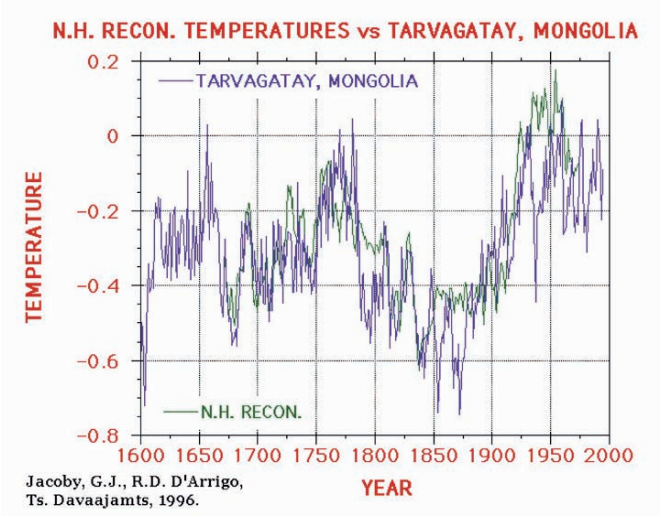


Figure 6.8: Temperatures derived from tree rings. Here the Maunder minimum is not very prominent whereas a cold period between 1830 to 1870)

products that are present. In the Pacific Ocean, transparency increases from east to west, with the greatest penetration of solar energy occurring in the western tropical Pacific. Due to ocean currents, the North Pacific Ocean takes approximately 4 years to move temperature anomalies from the western tropical Pacific to near North America (Favorite and McLain, 1973 [96]).

During the prolonged period between 1500 and 1850, average temperatures in Northern Europe were much colder than they are today, this is known as the little Ice Age. The coldest part of this period coincides with a conspicuous absence of sunspots and other signs of solar activity, called the Maunder Minimum.

For example Gilliland reported a 76 year cycle in the solar radius, inferred from a 258 year record of transits of the planet Mercury, solar eclipse records and meridian transit measurements. Ribes *et al.* (1987 [257]) also reported as Gilliland that the solar radius is slightly increased in times of low solar activity during the Maunder minimum.

A review book on the role of the Sun in climate change was written by Hoyt and Schatten (1997 [141]) where other references can be found.

Indicators of solar activity such as ^{14}C concentration measurements and of climate (e.g. glaciers) show a clear correlation over the last 7 000 years. This was shown by Eddy in 1977. Considering a time series since 1860 the solar cycle length shows an excellent correlation with northern hemisphere land temperatures (Friis-Christensen and Lassen, 1991 [103]). For recent data however, these two parameters diverge. If there exists a global climate contribution of solar irradiance variations then there are three possible interactions or couplings between these variations and the Earth's climate:

- Variations of the total solar irradiance.
- Variations of the Sun's spectral irradiance; this denotes changes in the luminosity of the Sun in a given wavelength range. As we have discussed above, UV radiation influences atmospheric chemistry (production or destruction of ozone, see also Haigh 1994 [122], 1996 [123]).
- Variations in the heliospheric magnetic field which are coupled to changes in the solar wind and influences the number and energy spectrum of cosmic ray particles. This was investigated by Potgieter, 1998 [247] and Simpson, 1998 [285]. The variation of cosmic ray particles seems to be related to global cloud cover (see Svensmark and Friis-Christensen (1997 [306]) or Svensmark (1998 [305]) or Marsch and Svensmark (2000 [211]).

As we have mentioned above, the total solar irradiance varies by 0.1% and these measurements have been made very accurately since 1978 (a review about that was given by Fröhlich, 2000 [102]) The cycles covered by these measurements are 21, 22 and 23. Of course from these time series it is impossible to extrapolate to earlier time series when the Sun was more active (e.g. cycle 19) or less active. One further problem of the time series available is that with the exception of SOHO/VIRGO they are restricted to the UV.

Irradiance variations of the past solar cycles can be determined from the surface distribution of the magnetic field if records of the field distribution or of proxies are available. The following proxies can be used:

- relative sunspot number (since 1700),
- group sunspot number (since 1610),
- sunspot and facular areas (A_s, A_f, since 1874,)
- Ca II plage areas (A_p, since 1915).

Using these data, one can reconstruct the cyclic component of the irradiance back to the Maunder minimum.

As a large sunspot group passes across the solar surface, there is a dip in the total solar irradiance. The variation is in the range of 0.02%.

Lockwood *et al.* (1999 [200]) reconstructed the aa-index of geomagnetic activity and found that the interplanetary magnetic flux at minimum of solar activity (that can be be reconstructed using the aa-index) has roughly doubled since 1900. This is in good agreement with ^{10}Be concentration in Greenland ice (Beer, 2000 [32]). ^{10}Be is produced by the interaction of cosmic rays with constituents of the Earth's atmosphere. The cosmic ray flux is modulated by the heliospheric magnetic field.

Lean *et al.* (1995 [191]) assumed that the background irradiance is proportional to the amplitude of the solar cycle; Hoyt and Schatten (1993 [140]) propose a trend corresponding to cycle length and Baliunas and Soon (1995 [25]) demonstrated that the amplitudes of stellar cycles (observed in Ca II H and K) scale with the length of the stellar cycle.

A short overview of long term chances in solar irradiance was given by Solanki and Fligge (2000 [293]).

The question whether the Earth's climate is influenced by solar activity has a central position in the present debate about the global warming. Greenhouse gas concentrations have a continuous increase and do not follow the observed decrease in the 1900's and in 1940-1970 example. These variations might be better explained when solar activity is taken into account. During a normal sunspot cycle the irradiance changes by 0.1% but could be greater (e.g. during the Maunder Minimum 0.3%, Lean (1997 [189])).

6.3.3 Solar Protons

Solar protons when hitting the atmosphere, break up molecules of N_2 and H_2O vapor. When nitrogen gas molecules split apart, they can create molecules, called nitrogen oxides NO, which can last several weeks to months depending on where they end up in the atmosphere. Once formed, the nitrogen oxides react quickly with ozone and reduce its amounts. When atmospheric winds blow them down into the middle stratosphere, they can stay there for months, and continue to keep ozone at a reduced level.

Similarly water vapor molecules are affected by solar protons, breaking them up into radicals where they react with ozone. However, these molecules, called hydrogen oxides, only last during the time period of the solar proton event. These short-term effects of hydrogen oxides can destroy up to 70 percent of the ozone in the middle mesosphere. At the same time, longer-term ozone loss caused by nitrogen oxides destroys a maximum of about nine percent of the ozone in the upper stratosphere. Only a few percent of total ozone is in the mesosphere and upper stratosphere with over 80 percent in the middle and lower stratosphere.

The impacts on humans are minimal. NASA's HALOE was launched on the UARS spacecraft September 15, 1991 as part of the Earth Science Enterprise Program. Its mission includes improvement of understanding stratospheric ozone depletion by measuring vertical profiles of ozone, hydrogen chloride, hydrogen fluoride, methane, water vapor, nitric oxide, nitrogen dioxide, aerosols, and temperature. The SBUV/2 instrument was launched aboard the NOAA-14 satellite on December 30, 1994 and its mission is to observe the ozone layer.

6.4 Cosmic Rays

6.4.1 Origination of Cosmic Rays

Victor Hess discovered in 1912 during a balloon flight to an altitude of more than 5000 m that the ionization rate increased. This fact he explained by the assumption that a radiation of very great energy penetrating power enters from above the atmosphere.[3] Energetic particles originating beyond the Earth that impinge on our atmosphere are called cosmic rays. The particles span energies

[3]Hess received the Nobel Prize in Physics in 1936

over a wide range and more than 80% are protons more than 12% He nuclei (α particles), the rest electrons, gamma rays and neutrinos. Because the particles are charged we have to take into account the interaction with magnetic fields, mainly with the heliosphere and the Earth's magnetic field. The particles are high energetic, thus they produce showers of particles when they collide with particles in our atmosphere (pions, kaons, mesons, muons).

There exist three components of cosmic rays:

- Galactic cosmic rays, GCR
- Anomalous cosmic rays, ACR
- Solar Energetic particles, SEP

Galactic Cosmic Rays

Since the particles are charged, they are deflected by the magnetic field of our Galaxy and the heliosphere as well as the Earth's magnetic field. Thus we can no longer point back to their sources in the Galaxy. A map of the sky with cosmic ray intensities would be completely uniform. The composition of the GCR's could tells us something about the origin. One finds all natural elements in roughly the same proportion as they occur in the solar system. Two properties of the particles can be measured:

- Determine which element; this is very easy since the different charges of each nucleus give different signatures.
- Isotopic composition; to determine the isotopic composition which in some cases gives better insights in the origin of the particles, the atomic nuclei have to be weighted which is much more difficult.

Most GCR are accelerated in the blast waves of supernova remnants. Remnants of supernova explosions are very active, we observe expanding clouds of gas, magnetic field activities etc. which can accelerate particles. Such processes in supernova remnants can last for several 10^3 yrs. The particles are accelerated in the magnetic field until they have enough energy to escape and become cosmic rays. Thus they can only be accelerated up to a certain maximum energy which depends upon the size of the acceleration region and the magnetic field strength.

The problem we face is that cosmic rays have been observed at much higher energies than those supernova remnants can generate. Possible explanations for such extreme high energetic particles are:

- their nature is extragalactic: from galaxies with very active galactic nuclei,
- they are related to the gamma ray bursts (energies up to 800 GeV are measured, they produce fast electrons and positrons in the Earth's atmosphere; in vacuum no particle can move faster than light. Air has a refractive index, therefore the speed of light is slowed down and fast enough electrons travel faster than light. Then, like the sonic shock ahead of a plane moving with supersonic speed they emit a shock front of light – Cherenkov radiation[4].

[4]You can see this radiation as the glow coming from a research reactor

Flashes of Cherenkov light from air showers have been studied for many years.

- they are related to exotic particles which are predicted by several physical theories concerning the origin of the universe; superstrings, exotic matter, strongly interacting neutrinos,

- they are topological defects in the very structure of the universe.

As we have stated above, cosmic rays include a number of radioactive nuclei whose numbers decrease through the radioactive decay. Measurements of these nuclei can be used therefore (as in the C^{14} method) to determine how long it has been since cosmic ray material was synthesized in the galactic magnetic field.

The origin of GCRs was discussed by Axford, 1994 [18].

One example of modern cosmic ray detectors is the Pierre Auger observatory installed in Argentina in 2005. 1600 water tanks, each containing 3000 gallons of water and separated by 1,5 km from each of its neighbors. The tanks are completely dark inside. Therefore we can measure the Cherenkov radiation of particles travelling faster than light. Slight differences in the detection times at different tank positions help scientists determine the trajectory of the incoming cosmic ray. The charged particles in an air shower also interact with atmospheric nitrogen, causing it to emit ultraviolet light via a process called fluorescence. The observatory's second detection method uses these detectors to observe the trail of nitrogen fluorescence.

Anomalous Cosmic Rays, ACR

The second component of cosmic ray particles originates from interaction with the heliosphere with neutral interstellar gas. Interstellar neutral gas flows through the solar system, since uncharged particles are not influenced by magnetic fields. The speed is approximately 25 km/s. When approaching the Sun, these neutral atoms become ionized by two processes:

- photo-ionization: an electron of the neutral atom is knocked off by a solar high energy photon (e.g. a UV photon);

- charge exchange: an electron is exchanged to an ionized atom of solar wind particle.

As soon as these particles are charged the Sun's magnetic field carries them outward to the solar wind termination shock region. The ions repeatedly collide with the termination shock, gaining energy during each collision. This continues until they escape from the shock region and diffuse back toward the inner heliosphere. Such particles are called anomalous cosmic rays (ACRs). ACRs are thought to originate from the very local interstellar medium and are not related to the above mentioned violent processes as the GCRs. They can easily be discerned from GCRs because they have lower speed and energy. They include large quantities of He, O, Ne and other elements which have in common high ionization potentials.

Figure 6.9: A collision between a high-energy cosmic ray particle and an atom in a photographic emulsion (viewed through a microscope). Below: filling of a water tank of the Auger observatory.

Solar Energetic Particles, SEP

The third component of cosmic ray particles are Solar energetic particles (*SEP*). They move away from the Sun due to plasma heating, acceleration and other processes. Flares e.g. inject large amounts of energetic nuclei into space, the composition varies from flare to flare. On the scale of cosmic radiation, SEPs have relatively low energies.

6.4.2 The Heliosphere

The heliosphere as already stated, is the magnetic shield caused by the Sun which protects us against energetic cosmic ray particles. The solar wind which is a continuous stream of plasma expands out through the solar system until it changes from supersonic to subsonic speed what is called a termination shock. The distance of that region is assumed to be about 90 AU and the space within is called the Heliosphere which encloses the whole solar system (e.g. Pluto's mean distance is only 39 AU). The termination shock was crossed by the spacecraft Voyager 1 on December 16, 2004 (see Stone *et al.*, 2005 [301]). The distance of the shock was at 84.01 AU from the Sun. It was found that this shock is a steady source of low energy protons. The intensity of anomalous cosmic ray helium did not peak at the shock, indicating that the ACR source is not in the shock region local to Voyager 1. The observations showed that the physics might be more complicated there. Voyager 2 is expected to cross the termination shock in 2008.

Because of the magnetic fields, only some of the GCR particles penetrate to the inner part of the solar system. Thus the magnetic field of the heliosphere works as a shield. The magnetic activity of the Sun changes however with the solar cycle (every 11 year the Sun's magnetic field reverses, the true cycle is thus 22 years). This causes a variation of the GCR flux. When the Sun is more active, the magnetic field is stronger, and as a result, fewer GCR arrive in the vicinity of the Earth. We can also say that the higher the energy the particles have, the less they are modulated by the solar cycle.

For our study here, it is important to note that by measuring cosmic rays one can derive a proxy for solar activity very long back in time. This is possible since isotopes in the atmosphere are produced by cosmic rays. From such recordings a good qualitative agreement between cold and warm climatic periods and low and high solar activity during the last 10 000 years was found. When we consider ^{14}C variations during the last millennium, one can deduce, that from 1000-1300 AC solar activity was very high which coincided with the warm medieval period. We know from history that e.g. during that period the Vikings settled in Greenland. The solar activity - if it is well represented by the ^{14}C variation- decreased and a long period followed which is now called the little ice age (in this period falls also the so called Maunder Minimum, 1645-1715, during which practically no sunspots were observed). This period lasted until the middle of the 19th century. From then on, solar activity has increased and is the highest in the last 600 years.

Thus we may assume the following connections:

$$\text{low solar activity} \rightarrow \text{weak magnetic field} \rightarrow \text{more GCRs} \rightarrow \text{more } ^{14}\text{C} \qquad (6.21)$$

If that assumption is true, there is a mechanism, how the Earth's climate can be influenced by the Sun.

The correlation between cosmic rays and solar activity and temperatures on Earth was studied by Usoskin *et al.*, 2005 [322]. Comparison of the Sun-related data sets with various reconstructions of terrestrial Northern Hemisphere mean surface temperatures reveals consistently positive correlation coefficients for the sunspot numbers and consistently negative correlation coefficients for the cosmic rays. The significance levels reach up to 99% but vary strongly for the different data sets.

Predictions of Galactic Cosmic Ray Intensity Deduced from that of Sunspot Number were made by Lantos, 2005 [188].

The relationship between cosmic ray variability and enhanced geomagnetic activity was summarized by Kudela and Storini, 2006 [178].

6.4.3 Clouds, Cloud Formation Processes

Clouds are created by condensation or deposition of water above the Earth's surface when the air mass becomes saturated (relative humidity 100 %). Saturation can occur when the temperature of an air mass goes to its dew point or frost point. There are different mechanisms to achieve this:

Orographic precipitation: this occurs when air is forced to rise because of the physical presence of elevation (land). As such a parcel of air rises it cools due to the adiabatic expansion at a rate of approximately 10 degrees per 1 000 m. The rise of the parcel is stopped if saturation is reached. An example of this mechanism is the west coast of Canada with large precipitation.

Convectional precipitation: this is associated with heating of the air at the Earth's surface. When there is enough heating, the air becomes lighter than the surrounding masses, begins to rise (cf. a hot air balloon begins to rise), expands and cools as above. When sufficient cooling takes place, saturation is reached again forming precipitation. This mechanism is active in the interior of continents and near the equator forming cumulus clouds and thunderstorms.

Convergence or frontal precipitation: this mechanism takes place when two masses of air come together. One is usually moist and warm and the other is cold and dry. The leading edge of the cold front acts as an inclined wall or front causing the moist warm air to be lifted. Then the above described processes start again: rise, cooling and saturation. This type of precipitation is common in the mid latitudes.

Finally we have to mention the radiative cooling: this occurs when the Sun is no longer supplying the ground and overlying air with energy due to insolation during nighttime. The surface of the Earth begins to lose energy in the form of longwave radiation. This causes the ground and the air above it to cool. The precipitation that results from this kind of mechanism takes the form of dew, frost or fog.

Of course these mechanisms may act as a combination: convection and orographic uplift can cause summer afternoon showers in the mountains.

Let us compare the levels of cloud cover for summer and winter (northern hemisphere). For summer in the northern hemisphere, highest levels of cloud cover occur over the mid-latitude cyclone storm tracks of both hemispheres, Intertropical Convergence Zone over land surfaces, and the Indian Monsoon region (orographic lifting). Lowest values occur over the subtropical deserts, the subsidence regions of the subtropical oceans, and the polar regions. For winter in the northern hemisphere highest levels of cloud cover occur over the mid-latitude cyclone storm tracks of both hemispheres and the Intertropical Convergence Zone over land surfaces. Lowest values occur over the subtropical deserts, the subsidence regions of the subtropical oceans, and over the South Pole.

Clouds influence vertically integrated radiative properties of the atmosphere. They cause a cooling through reflection of incoming shortwave radiation (sun - light) and heating by absorption and trapping of outgoing long wave radiation (thermal radiation). Let us consider the net radiative impact of a cloud: this mainly depends on two parameters, on its height above the surface and its optical thickness. High optically thin clouds tend to heat while low optically thick clouds tend to cool. The net forcing of the global cloud cover is in the range between $17 - 35\,\mathrm{W/m^2}$, as it is derived from climate models. Thus a significant influence on the global cloud cover can be potentially very important for Earth's climate (see also Table 6.5).

It has been found that the Earth's cloud cover follows the variation in GCR.

Table 6.5: Various influences on the climate

S, solar constant (at 1 AU)	1360 W/m^2
S/4, top of atmosphere	340 W/m^2
S/4(1-a), a=0.3 Earth's albedo	235 W/m^2
1 % change in the albedo	1 W/m^2
estimated rad. effect due to CO_2 increase since 1750	1.5 W/m^2
doubling of CO_2	4 W/m^2
radiative effect of clouds (cooling)	17-35 W/m^2

It seems to be that the ionization in the atmosphere produced by GCR is the essential link. One can estimate that a variation in cloud cover of 3 % during an average 11-year solar cycle could have an effect of 0.8-1.7 W/m^2. This is a very significant amount.

The idea that cosmic rays can influence cloud formation was first pointed out by Svensmark *et al.* (1997 [306]). They showed that there was a significant correlation between total cloud cover over the Earth and the influx of cosmic rays. The rays ionize particles in the low troposphere which then seed the growth of cloud water droplets. During the past century the shielding from cosmic rays has increased since solar activity has increased. This decreases the formation and cooling influence of low clouds and may thus provide a possible contribution to the global warming of the past 100 years (Marsh, Svensmark, 2000 [212]).

Let us consider and summarize the changes in the magnetic field in the solar atmosphere. Shorter solar cycles facilitate a rise in the coronal source flux, longer cycles allow it to decay. The accumulation of the coronal source flux strongly depends on the rate of flux emergence in active regions. In general the peak and cycle averaged sunspot numbers are larger when cycles are shorter. Therefore, shorter cycles are associated with larger flux emergence rates, there is less time for the open flux to decay. We can state:

shorter activity cycle $\rightarrow$ increased coronal flux

The coronal source surface is where the magnetic field becomes approximately radial. This occurs at $r = 2.5R_{\odot}$. This surface can also be regarded as the boundary that separates the solar corona from the heliosphere. The magnetic flux threading the corona source surface is called F_s or open solar flux. If there is a rise of the flux F_s than the cosmic ray flux incident on the Earth will decrease. Lockwood and Foster (2000 [199]) estimated that the cosmic ray flux > 3 GeV was 15% larger around 1900 than it is now. As it was shown above, cosmic rays generate air ions in the sub ionospheric gap which allows current to flow in the global electric current. This connects thunderclouds with the ground via lightening.

An analysis of ISCCP D2 cloud data showed a correspondence between low cloud cover and cosmic ray flux (Palle and Butler, 2000 [242]). The authors also mentioned that the effect of increased global sea temperatures, increased aerosols and aircraft traffic on cloud formation processes should be taken into account.

Table 6.6: Causes of Global Warming of about 0.5 C, 1880-1997

Climate Forcing Factor	Est. forcing °C , 1880-1997
Solar luminosity increase	+0.25[1]
Decrease in volcanic stratospheric aerosols	+0.15[2]
Increased anthropogenic sulfate aerosols	Up to -0.1 C
Increased anthropogenic carbon aerosols	Up to + 0.1 C (offsets sulfate aerosols)
Carbon dioxide warming	+0.05 to +0.10 C
Decrease in stratospheric ozone	-0.05 C[3]
Increase in cirrus contrails from airplanes	+ 0.05 C
Urban heat island effects	+0.01 to +0.10[4]
Changing skyline effects	possibly as large as + 0.25 C
Sum total of all above forcing factor	+0.51 to 0.60 C

[1] See "The Role of the Sun in Climate Change"; also see Lean *et al.*, 1995 [191].
[2] Wu *et al.*, 1999 [341]
[3] Schwartz and Andreae 1996
[4] Balling, 1992

6.5 What Causes the Global Warming?

This is a very strong debate. In the extreme case the warming is not caused by a substantial greenhouse effect since as is shown in Table 6.6 other factors can contribute to the observed increase in temperature. However, all these estimates are estimates and should be taken with caution.

In Fig 6.10 a graph from the summary for policymakers of the report of WG 1 of the intergovernmental panel on climate change illustrates the estimated global mean radiative forcing of the climate system for the year 2000 relative to 1750. The main forcing due to CFC's is easily recognized, the influence of the Sun and its variation is marked as bad known but as more important as e.g. black carbon from fossil fuel burning.

Generally, the effect of cosmic rays seems to be important but cannot explain the whole climate warming observed during the past 100 years. The climate response to changes of cosmic ray flux was investigated by Shaviv, 2005[281]. The result can be summarized as:

- there is certainly a link between the cosmic ray flux (CRF) and solar luminosity
- the increased solar luminosity during the last 100 years lead to a decreased CRF

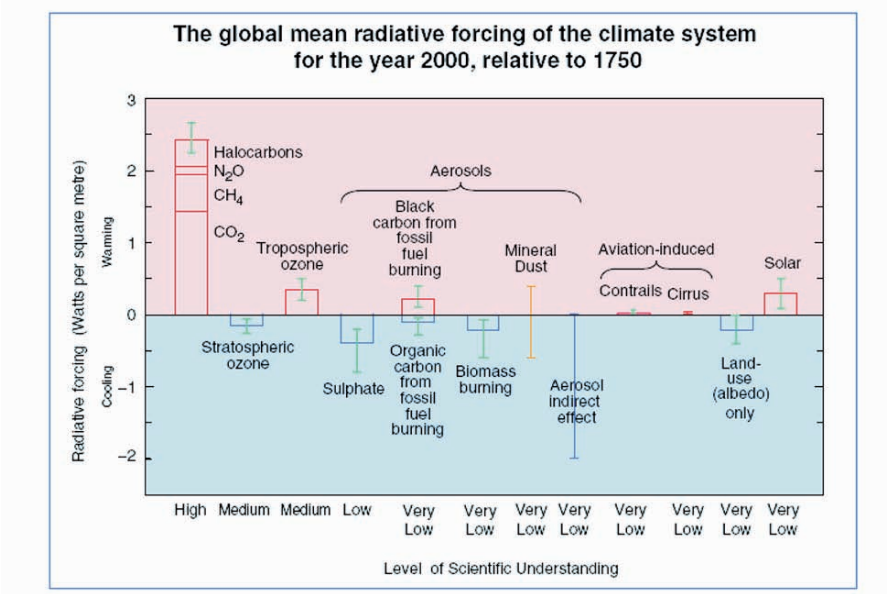


Figure 6.10: Global radiative cooling and warming of the climate system (from IPCC report).

- CRF/climate link therefore implies that the increased solar luminosity and reduced CRF over the previous century should have contributed a warming of 0.47 ± 0.19 K,
- the rest should be mainly attributed to anthropogenic causes.
- Without any effect of cosmic rays, the increase in solar luminosity would correspond to an increased temperature of 0.16 ± 0.04 K

Meteorite data on the galactic cosmic rays, the solar activity, and temperature variations in the earth's atmosphere lead to the conclusion that the solar activity may be important factor exerting the influence upon the climate of the Earth (see e.g. Alexeev, Ustinova, 2005[4]). Estimations on the long term cosmic ray variation and possible climate on planets were made by Dorman, 2005 [78].

A summary of the effects is illustrated in Fig. 6.11. The data shown in that Fig are i) Reconstructed NH temperature series from 1610-1980, updated with raw data from 1981-1995 ii) Greenhouse gases (GHG) represented by atmospheric CO2 measurements (iii) Reconstructed solar irradiance (see Lean *et al.*, 1995 [191]) (iv) Weighted volcanic dust veil index (DVI) (v) Evolving multivariate correlation of

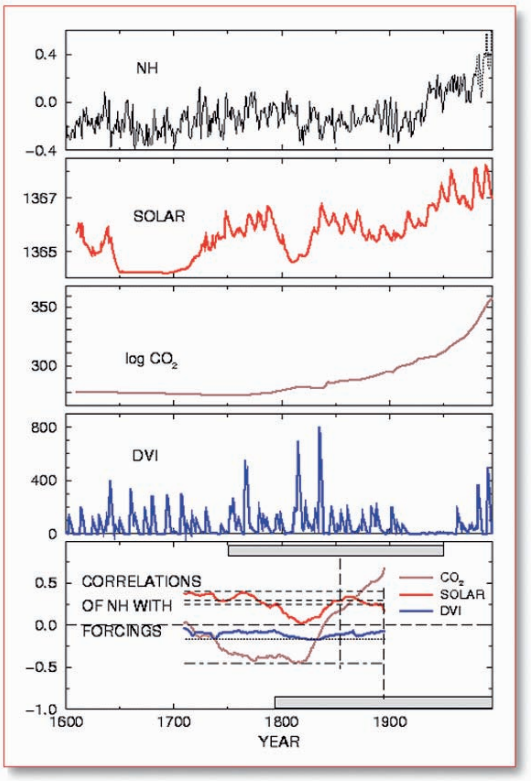


Figure 6.11: Relationship of Northern hemisphere mean (NH) temperature reconstruction to estimates of three candidate forcings between 1610 and 1995.

NH series with the 3 forcings (i) (ii) and (iii). The data are from Mann *et al.* (1999 [206], and further references therein). These authors conclude that while the natural (solar and volcanic) forcings appear to be important factors governing the natural variations of temperatures in past centuries, only human greenhouse gas forcing alone, can statistically explain the unusual warmth of the past few decades.

Chapter 7

Space Weather and Radiation Damage

In this chapter we discuss the influences of radiation damage both to humans in space as well as to electronics and solar panels of satellites. A general overview on the radiation environment in the interplanetary space was given by Townsend and Wilson, 1996 [315].

7.1 Radiation Damage on Living Organisms

7.1.1 Definitions

Radiation is energy in the form of waves or particles. X rays and gamma rays are electromagnetic waves of radiation, as is visible light. Particulate radiation includes alpha and beta radiation. The energy associated with any radiation can be transferred to matter. This transfer of energy can remove electrons from the atoms leading to the formation of ions. The types of radiation capable of producing ions in matter are collectively referred to as *ionizing radiation*.

Alpha particles are composed of two protons and two neutrons. Alpha particles do not travel very far from their radioactive source. They cannot pass through a piece of paper, clothes, or even the layer of dead cells which normally protects the skin. Because alpha particles cannot penetrate human skin they are not considered an *external exposure hazard* (this means that if the alpha particles stay outside the human body they cannot harm it). However, alpha particle sources located within the body may pose an "internal" health hazard if they are present in great enough quantities. The risk from indoor radon is due to inhaled alpha particle sources which irradiate lung tissue.

Beta particles are electrons not bound to any atom. Beta particles cannot travel very far from their radioactive source. For example, they can travel only about one cm in human tissue, and they may travel a m in air. They are not capable of penetrating something as thin as a book or a pad of paper.

Table 7.1: Radiation related units

Unit	Measures	Definition
Roengten (R)	exposure	1 R=2.56×10^{-4} C/s is deposited in dry air kg^{-1}; only for X rays
Radiation absorbed dose (rad)	absorbed dose	1 rad = absorption of 100 ergs per g material used for any type of radiation no description of biol. effects
Roengten equivalent man REM	equivalent dose	rem=rad × Q Q... quality factor (type of radiation) relates absorbed dose to effective biological damage
Gray, Gy	absorbed dose	1 Gy= 1J of energy deposed in 1 kg of material
Sievert Sv	equivalent dose	Sv = Gy × Q, Q...quality factor 1 Sv = 100 rem
Becquerel Bq	radioactivity	1Bq=1 transformation/sec 1 Cu=3.7×10^{10}Bq

Gamma rays are an example of electromagnetic radiation, as is visible light. Gamma rays originate from the nucleus of an atom by nuclear transitions. They are capable of travelling long distances through air and most other materials. Gamma rays require more "shielding" material, such as lead or steel, to reduce their numbers than is required for alpha and beta particles.

In Table 7.1 we give some definitions used in radiation physics.

The effect of radiation on any material is determined by the *dose* of radiation that material receives. Radiation dose is simply the quantity of radiation energy deposited in a material. There are several terms used in radiation protection to precisely describe the various aspects associated with the concept of dose and how radiation energy deposited in tissue affects humans.

Some terms related to radiation dose:

- Chronic dose: A chronic dose means a person received a radiation dose over a long period of time.
- Acute dose: An acute dose means a person received a radiation dose over a short period of time.
- Somatic effects are effects from some agent, like radiation that are seen in the individual who receives the agent.
- Genetic effects: Genetic effects are effects from some agent, that are seen in the offspring of the individual who received the agent. The agent must be encountered pre-conception.

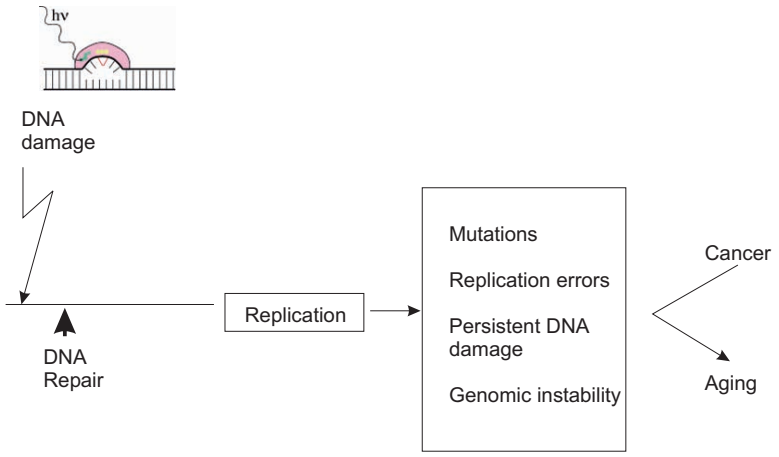


Figure 7.1: DNA damage caused by radiation

- Teratogenic effects: Teratogenic effects are effects from some agent, that are seen in the offspring of the individual who received the agent. The agent must be encountered during the gestation period.

7.1.2 Radiation Damage on DNA

The basic unit of any living organism is a cell. It is a small, watery compartment filled with chemicals and a complete copy of the organism's genome. The term genome denotes all the DNA in the cell (chromosomes and other). Different organisms have different numbers of chromosomes (e.g. humans have 23 pairs of chromosomes, 44 autosomes and 2 pairs of sex chromosomes). Each parent contributes one chromosome to each pair and so children get half of their chromosomes form their mother and half from their father.

The structural arrangement of DNA looks like a long ladder twisted into a helix. The sides of the ladder are formed by a backbone of sugar and phosphate molecules. The rungs consist of nucleotide bases joined weakly in the middle by hydrogen bounds. There are two major ways that radiation injures the DNA inside the cells of an organism:

- Water in the body tends to absorb a large fraction of radiation and becomes ionized. When water is ionized it forms highly reactive molecules which are called free radicals. Those react with and damage the DNA molecules.

- Radiation can also collide directly with the DNA molecules ionizing and damaging it directly.

The typical symptoms of radiation sickness are: severe burns that are slow to heal, sterilization, cancer. High doses are rapidly fatal (within days or weeks). Yang *et al.* (1996 [342]) discussed DNA damage and repair in oncogenic transformation

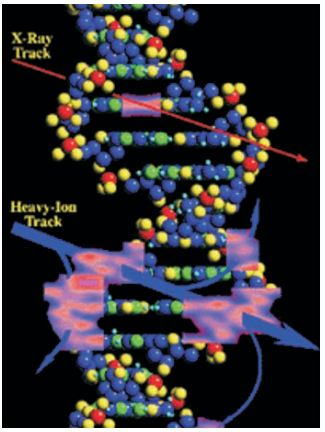


Figure 7.2: Passage of ionizing radiation can result in direct effect on DNA leading to single strand breaks (SSB), double strand breaks (DSB), associated base damage (BD), or clusters of these damage types. Source: NASA

by heavy ion radiation. The most important late effect of energetic heavy ions in cosmic rays and solar particle events is risk assessment in carcinogenesis.

7.1.3 DNA Repair

Wether or not a cell can repair depends on the damage to the DNA.

- single strand break in the DNA: this can be usually repaired and normal cell function is restored.
- breaks in both DNA strands: usually the damage is too severe to repair and the cell dies.
- chemical change or mutation: cannot be repaired; cancer or a mutation offspring results if this occurs in a sperm or egg cell.

7.1.4 Radiation Dose Limits for Astronauts

These limits were set by the US National Council on Radiation Protection and Measurements for all space missions in order to protect the astronauts. But there is an exception for exploratory missions and circumstances in space (e.g. mission to Mars). In Table 7.2 the relevant data are given.

The radiation dose limits for ordinary citizens are much lower (see Table 7.3). The annual dose is about 50 mSv, the lifetime dose is age [years] x 10 mSv. In the US the total average annual dose is about 3.6 mSv.

The single dose effects are described in Table 7.4

Riklis *et al.* (1996 [260]) discussed biochemical radioprotection using antioxidants and DNA repair enhancement and found that the right combination proves

Table 7.2: Radiation dose limits in mSv for astronauts

Time period	blood forming organs	eyes	skin
30 days	250	1000	1500
annual	500	2000	3000
career for males	2000 mSv+75(age[years]-30)	4000	6000
career for females	2000 mSv+75(age[years]-38)	4000	6000

Table 7.3: Total average annual radiation does in the US

radon in the air	2 mSv (56%)
rocks, building material	0.28 mSv (8%)
cosmic rays	0.28 mSv (8%)
natural radioactive material in body	0.39mSv (11%)
medical and dental rays	0.39 mSv (11%)
nuclear medicine tests	0.14 mSv (4%)

Table 7.4: Single dose effects

0.001 mSv	dental x rays
0.002 mSv	5 hr transcontinental flight
0.02 mSv	chest X-ray
1.000 mSv	radiation sickness
2500 mSv	sterility in females
3500 mSv	sterility in males
4000 mSv	average lethal dose (without any treatment)

effective in providing protection from a wide range of radiation exposures over a long period of time.

7.1.5 Genetic vs. Somatic Effects

Somatic effects of radiation damage appear on the exposed person. Prompt somatic effects appear after an acute dose. One example of a prompt effect is temporary hair loss. Delayed somatic effects may occur years after radiation doses are received. Typical effects are the development of cancer and cataracts. Let us briefly mention the most important syndromes.

- Blood forming organ (bone marrow) syndrome: damage to the cells which divide at the most rapid pace; bone marrow, spleen and limphathic tissue. Symptoms include internal bleeding, fatigue, bacterial infections and fever.

- Gastrointestinal tract syndrome (>1000 rad): damage to cells which divide less rapidly; lining to the stomach and intestines. Symptoms are nausea, vomitting, diarrhoea, dehydration, loss of digestion ability, bleeding ulcers.

- Central nervous system syndrome (> 5000 rad): damage to cells which do not reproduce such as nerve cells. Symptoms include loss of coordination, confusion, coma, shock.

It seems now that death is not caused by radiation damage on the nervous system but by internal bleeding and fluid and pressure build-up on the brain.

The genetic or heritable effect appears in the future generation of the exposed person as a result of radiation damage to the reproductive cells.

We have seen that satellite systems are vulnerable to Space Weather through its influence on energetic charged particle and plasma populations and that aircraft electronics and air crew are subjected to atmospheric secondary radiation produced by cosmic rays and solar particle events. This is discussed by Dyer (2001 [85]). The Advanced Composition Explorer (ACE) continuously monitors the solar wind and produces warnings by monitoring the high-energy particles that can produce radiation damage in satellite systems (Zwickl *et al.*, 1998 [347]).

3.9-2.5 Billion years ago the Earth was dominated by an oceanic lithosphere. Cockell, 2000, [69] calculated that the DNA damage rates might have been approximately three orders of magnitude higher in the surface layer of the Archean oceans than on present-day oceans. However, at 30 m depth, damage might have been similar to the surface of present-day oceans. On the other hand, risk of being transported to the surface water in the mixed layer was quite high. Thus the mixed layer may have been inhabited by a low diversity UV-resistant biota. Repair capabilities similar to Deinococcus radiodurans would have been sufficient to survive in the mixed layer. During the early Proterozoic ozone concentrations increased and the UV stress would have been reduced and a greater diversity of organisms could have inhabited the mixed layer.

In STS Shuttle/Mir mission experiment, recovery of bacterial cells from radiation damage and the effects of microgravity were examined for Deinococcus radiodurans (Kobayashi *et al.*, 2000 [170]).

Lean (2000 [190]) discusses societal impacts of solar electromagnetic radiation. Climate change and ozone depletion has significant economic and political impacts on an international level.

The Yohkoh satellite was launched in 1991. Song and Cao (1999 [296]) discuss CCD radiation damage. Evans *et al.* (1999 [94]) discuss charged-particle induced radiation damage of a HPGe gamma-ray detector during spaceflight.

7.1.6 The Solar Proton Event in August 1972

Between the manned Apollo 16 and 17 missions one of the largest solar proton events ever recorded occurred. As a matter of luck no astronauts were in space during that time. Computer simulations were done later to reconstruct the influence on astronauts during that time. The main result of these simulations was that even inside of a spacecraft the astronauts would have absorbed a lethal dose

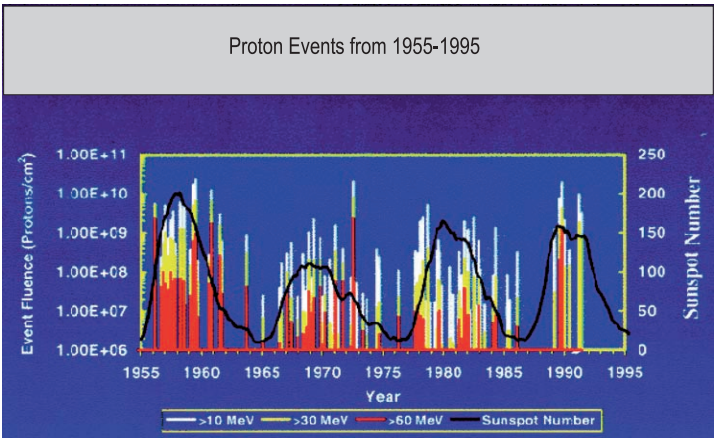


Figure 7.3: Correlation of the occurrence of solar proton events with solar activity cycle (indicated by the sunspot number)

of radiation within 10 hrs after the start of the event. At 6:20 UT an optical flare was observed on the Sun. At 13:00 UT the astronauts' allowable 30- day radiation exposure to skin and eyes was exceeded. At 14:00 the astronauts' allowable 30-day radiation exposure for blood forming organs and yearly limit for eyes was exceeded. The yearly limit for skin was exceeded at 15:00 UT. At 16:00 UT the yearly limit for blood forming organs and the career limit for eyes was exceeded. At 17:00 UT the career limit for skin was exceeded.

This event dramatically shows the need for space weather forecasting. The correlation of solar proton events with activity cycle is evident (Fig. 7.1.6).

Heckman (1988 [130]) discussed proton event predictions.

7.2 Solar UV Radiation Damage

7.2.1 General Remarks

Most UV radiation from the Sun is absorbed by the ozone layer or reflected back into space so only a small amount reaches the surface of the Earth. Sunlight is received as direct rays and as diffuse light, i.e. skylight which has been scattered by the atmosphere. The sky is blue because air molecules scatter the shorter wavelength (blue light) more than the red light, the index of scattering depends on the wavelength. UV light is scattered even more than blue light.

Due to diffuse UV light, being shaded from direct Sunlight provides only a partial protection. Typical window glasses transmit less than 10% of ultraviolet light, and sunblock creams work by absorbing or reflecting UV rays. The SPF rating of sunscreens gives an indication of their effectiveness as UV blockers. For example, an SPF of 15 means that it should take 15 times as long to before skin

damage occurs (i.e., the cream should block about 93% of the radiation that causes skin damage). UV radiation is subdivided into three wavelength bands:

- UVA (315-400 nm), produces photochemical smog; damages plastic, paints and fabrics. UVA rays are not as energetic as UVB and, as a consequence, cause little sunburn or skin reddening. On the other hand, UVA rays penetrate deeper into the skin. The damage they cause is on a cellular level, occurring slowly and accumulating over a period of time. UVA radiation induces the formation of free radicals that, in turn, attack the lipids in the skin. The resulting damage gives rise to the visible signs of aging such as wrinkles and thickened skin. The skin's natural defenses against these free radicals are ascorbic acid (vitamin C) and alpha-tocopherol (vitamin E). These two vitamins are potent anti-oxidants that intercept the free radicals before they can do much damage. Vitamin C protects significantly better against UVA phototoxicity than vitamin E. Vitamin E, on the other hand, is more efficient against UVB.

- UVB (290-315 nm); 1% of solar radiation energy is in this band, most of it absorbed by ozone. Can damage DNA; smaller changes in ozone can lead to large changes in UVB radiation at the surface. Other effects are: Production of vitamin D in humans, skin cancer and damage to eye tissue. Plants and aquatic organisms suffer reduced growth, and many materials such as plastics degrade more rapidly in response to increased UVB radiation.

- UVC (220-290 nm); totally blocked by ozone and other gases, does not reach the Earth's surface.

A person's potential to develop skin cancer is related to their exposure UVB radiation (sunburn). In New Zealand, about one person in three will develop a skin cancer during their lifetime. About half the number killed on the roads die of skin cancer in New Zealand. New Zealand and Australia have a very high melanoma incidence compared with other countries.

How can this be explained?

- New Zealanders have an outdoor lifestyle,

- wear fewer clothes now than in the past,

- the ancestors of most white-skinned New Zealanders migrated from the UK, which is at much higher latitude, and has much lower levels of UV radiation. These people are therefore poorly adapted to the relatively high levels of UV naturally present in New Zealand;

- calculations suggest that locations in the Southern Hemisphere should receive approximately 15% more UV than locations at a similar latitude north of the Equator (Basher, 1981 [27]). This is caused by differences in ozone between the Northern and Southern Hemispheres, and also because the Earth is slightly closer to the Sun during the Southern Hemisphere summer (McKenzie and Elwood, 1990 [220]);

- measurements show much larger differences, with biologically-damaging UV being 50-80% more in the Southern Hemisphere than at comparable Northern latitudes in Europe. The differences are caused by the buildup of tropospheric pollution (tropospheric ozone and aerosols) in the North (Seckmeyer and McKenzie, 1992 [277]);

- Much higher levels of UV are experienced in countries, such as Australia, which are closer to the equator.

The amount of UVB radiation at ground level is determined by three factors: a) solar elevation, b) the amount of ozone in the atmosphere and c) the cloudiness of the sky. Please note that during local noon the amount of background radiation is the same as direct radiation three hours before and afterwards. At NZ's latitude, approximately 40% of the daily sunburn radiation occurs during the two hour period centered on solar noon.

Since the late 1970s an ozone hole has formed over Antarctica during early spring. The amount of ozone over New Zealand varies seasonally with a maximum in spring and a minimum in early autumn. Evidence of ozone destruction has also been observed over the Arctic. The ozone hole is caused by the special meteorological conditions of the cold atmosphere above polar regions which amplify the destructive ability of CFCs. The Antarctic ozone hole cannot shift over New Zealand. However, ozone losses over Antarctica may contribute to changes in ozone over the whole globe. After emission, halogen source gases are either removed from the atmosphere or undergo chemical conversion. The time to remove about 2/3 of a gas is called the atmospheric lifetime. Thus the amount of a halogen source gas in the atmosphere depends on a) lifetime b) amount emitted to the atmosphere. The atmospheric lifetime of CFC-12 and CFC-113 is about 100 years.

In Fig. 7.4 clear sky UV indices for different stations in NZ are given. Seasonal variations are higher at low latitudes as well as the absolute values.

7.2.2 UV Radiation and Materials

Synthetic polymers such as plastics are widespread and used all over the world. Wood wich can be considered as naturally occurring polymer. Both are used in building construction and other outdoor applications. The UV-B content affects adversely the mechanical properties of these materials. Therefore, photostabilizers in the case of plastics and protective surface coatings in the case of wood have to be used. Increased UV radiation and increased temperature contribute to a reduction of service life of these materials. It is estimated that especially the developing countries will suffer from UV radiation damage. The ozone-layer depletion is occurs mainly at higher latitudes. But at these latitudes the temperatures are moderate and the degradation reactions of the above cited polymers is low. The change in the ozone column at low latitudes is small, but the ambient temperatures are high as well as a high solar UV-B radiation. Therefore the service life of plastics will be reduced. A well known effect of degradation is yellowing discoloration

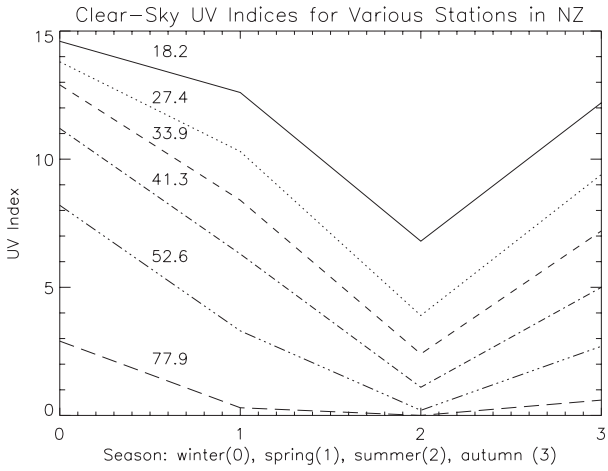


Figure 7.4: Typical clear-sky UV indices over New Zealand and its surrounding region. Seasonal variations are larger at low latitudes (denoted by numbers).

(e.g. wood, PVC) and loss of mechanical integrity (see also Andrady *et al.* 1998 [10]).

7.2.3 Effects on the Skin

Generally, excessive UV exposure results in a number of chronic skin changes. These include:

- various skin cancers of which melanoma is the most life-threatening;
- an increased number of moles (benign abnormalities of melanocytes),
- a range of other alterations arising from UV damage to keratinocytes and blood vessels;
- UV damage to fibrous tissue is often described as "photoageing". Photoageing makes people look older because their skin loses its tightness and so sags or wrinkles.

United Nations Environment Programme (UNEP) has estimated that more than 2 million non melanoma skin cancers and 200 000 malignant melanomas occur globally each year. Let us assume that there is a 10% decrease of stratospheric ozone; then it is estimated that an additional 300 000 nonmelanoma and 4 500 melanoma skin cancers would result worldwide.

The relationship between stratospheric ozone depletion and skin cancer was studied by many authors (see e.g. Amron and Moy, 1991 [8]). The impact of skin cancer related to climate change on the British population was reviewed by Diffey, 2004 [75].

Caucasians have a higher risk of skin cancer because of the relative lack of skin pigmentation. The worldwide incidence of malignant melanoma continues to increase, and is strongly related to frequency of recreational exposure to the sun and to history of sunburn. There is evidence that risk of melanoma is also related to intermittent exposure to UV, especially in childhood, and to exposure to sunlamps. However, the latter results are still preliminary.

7.2.4 Effects on the Eye

The acute effects of UV on the eye include the development of photokeratitis and photoconjunctivitis, which are like sunburn of the delicate skin-like tissue on the surface of the eyeball (cornea) and eyelids. While painful, they are reversible, easily prevented by protective eyewear and have not been associated with any long-term damage.

Chronic effects however include the possible development of pterygium (a white or cream colored opaque growth attached to the cornea), squamous cell cancer of the conjunctiva (scaly or plate-like malignancy) and cataracts. Some 20 million people worldwide are currently blind as a result of cataracts. Of these, WHO estimates that as many as 20% may be due to UV exposure. Experts believe that each 1% sustained decrease in stratospheric ozone would result in an increase of 0.5% in the number of cataracts caused by solar UV. Direct viewing of the sun and other extremely bright objects can also seriously damage the very sensitive part of the retina called the yellow spot, fovea or macula leutea. When cells of the fovea are destroyed, people can no longer view fine detail. For those people it becomes impossible to read, sew, watch TV, recognize faces, drive a vehicle etc.

7.2.5 Immune System

UV also appears to alter immune response by changing the activity and distribution of the cells responsible for triggering these responses. A number of studies indicate that UV exposures at environmental levels suppress immune responses in both rodents and humans. In rodents, this immune suppression results in enhanced susceptibility to certain infectious diseases with skin involvement, and some systemic infections. Mechanisms associated with UV-induced immunosuppression and host defence that protect against infectious agents are similar in rodents and humans. It is therefore reasonable to assume that UV exposure may enhance the risk of infection and decrease the effectiveness of vaccines in humans. Additional research is necessary to substantiate this.

7.2.6 UV Index

The Global Solar UV Index was developed through the WHO. It provides an estimate of the maximum solar UV exposure at the Earth's surface. The intensity of UV reaches a maximum around mid-day (when there is no cloud cover) at solar noon.

Table 7.5: The environment in space

Source	Energy	Hazard
Cosmic rays	GeV	SEU, Latchup
Solar flare particles	MeV to GeV	Interference
Radiation belt particles	MeV	Rad. damage, degradation
Energetic plasma	keV to MeV	Charging
Low energy plasma	eV to keV	leakage, sputtering
Neutral O atoms		Erosion
Debris		Puncture

It is generally presented as a forecast of the maximum amount of skin-damaging UV expected to reach the Earth's surface at solar noon. The values of the Index range from zero upward; the higher the Index number, the greater the likelihood of skin and eye damaging exposure to UV, and the less time it takes for damage to occur.

Close to the equator, summer-time values reach 20. During a European summer a value of 8 can be reached. We speak of:

- low UV exposure: Index 1...2
- moderate UV exposure: Index 3...4
- high UV exposure: Index 5...6
- very high UV exposure: Index 7...8
- extreme UV exposure: Index > 9

7.3 Radiation in Space

7.3.1 Space Environment

Outer space is extremely hostile and without a spacesuit :

- you would become unconscious within 15 s because there is no O,
- blood and other body fluids start to boil and then freeze because there is no air pressure,
- tissues (skin, heart...) expand because of the boiling fluids,
- extreme temperature changes: sunlight 120^0C, shade -100^0 C.
- exposure to radiation and micrometeoroids.

In Table 7.5 we summarize the different environmental effects in space.

The space environment is extremely hostile and protection must be provided.

The ISS station is at a height of about 400 km. There is some protection from the Earth's magnetosphere concerning charged particles. The Astronauts that travelled to the moon absorbed higher doses, up to 3 times.

For the planned manned mission to Mars for which a duration of 1000 days is proposed it is estimated that under the very best protection the risk for the Astronauts dying from cancer could be doubled.

7.3.2 The Extravehicular Mobility Unit

Some facts: Weight = 127 kg on Earth, Thickness = 0.48 cm, 13 layers, atmosphere = 0.29 atm of pure oxygen, Volume = 0.125 to 0.153 m^3, without astronaut cost = 12 million USD.

While early spacesuits were made entirely of soft fabrics, the EMU has a combination of soft and hard components to provide support, mobility and comfort. The suit itself has 13 layers of material, including an inner cooling garment (two layers), pressure garment (two layers), thermal micrometeorid garment (eight layers) and outer cover (one layer). The materials used include: Nylon tricot Spandex, Urethane-coated Nylon, Dacron, Neoprene-coated Nylon, Mylar, Gortex, Kevlar (material in bullet-proof vests), Nomex. All of the layers are sewn and cemented together to form the suit. In contrast to early spacesuits, which were individually tailored for each astronaut, the EMU has component pieces of varying sizes that can be put together to fit any given astronaut. The EMU consists of the following parts:

- Maximum Absorption Garment (MAG) - collects urine produced by the astronaut. Liquid Cooling and Ventilation Garment (LCVG) - removes excess body heat produced by the astronaut during spacewalks EMU.

- Electrical Harness (EEH) - provides connections for communications and bio-instruments.

- Communications Carrier Assembly (CCA) - contains microphones and earphones for communications.

- Lower Torso Assembly (LTA) - lower half of the EMU including pants, knee and ankle joints, boots and lower waist Hard Upper Torso (HUT) - hard fiberglass shell that supports several structures including the arms, torso, helmet, life-support backpack and control module Arms Gloves - outer and inner gloves Helmet.

- Extravehicular Visor Assembly (EVA) - protects the astronaut from bright sunlight

- In-suit Drink Bag (IDB) - provides drinking water for the astronaut during the spacewalk.

- Primary Life Support Subsystem (PLSS) - provides oxygen, power, carbon dioxide removal, cooling water, radio equipment and warning system.
- Secondary Oxygen Pack (SOP) - provides emergency oxygen supply.
- Display and Control Module (DCM) - displays and controls to run the PLSS

7.3.3 Radiation Shielding

Since the 1950s it is known that radiation in space poses a problem to human space travel. In 1952 Wernher von Braun and other space visionaries suggested using lunar soil to protect manned expedition from space radiation and meteors.

Low energy radiation can be stopped by a spacecraft wall. At higher energies the wall itself produces showers of secondary radiation and even more shielding is needed to absorb that. Using light weight materials like hydrogen, boron and lithium, nuclei of heavy elements in cosmic rays can be shattered by lightweight atoms without producing additional hazardous recoil products like neutrons. Thus, composites and other materials using low mass atoms might provide good shielding. At NASA's Langley Research Center simulated Martian soil will be tested for shielding.

The International Space Station (ISS) at 51.6^0 inclination and 220 mile of altitude is being constructed during a period of high solar activity with about 1000 hours of required extra vehicular activity (EVA). The Astronauts are exposed to trapped protons and electrons and galactic cosmic rays. Especially during transits through the South Atlantic Anomaly (SAA) EVA astronauts may experience enhanced doses. Dose enhancements are also expected from solar particle events (SPE). There are two different types of suits for astronauts: EMU and Orlan.

The skin responses to radiation include erythema, epilation, desquamation. Different anatomical skin sites vary in sensitivity with decreased order of responsiveness as follows:

- anterior aspect of neck,
- anterior surfaces of extremities, chest, abdomen,
- face
- back, posterior surfaces,
- nape of neck,
- scalp, palms, soles.

Literature: For more details on the problem see e.g. Kiefer (2001 [161]) or the Space Studies Board of the National Research Council (2000) or Thomson (1999 [312]), Badhwar (1997 [20]). Radiation measurements on Russian spacecraft Mir are presented by MacKay et al. (1993 [219]).

The effectiveness of any shielding depends on the energy distribution of the incident radiation. Some examples for common shielding materials are listed in Table 7.6.

Table 7.6: Common shielding materials

lead	11.35 g/cm^3
aluminium	2.7
water	1.0
lithium hydride	0.82
liquid hydrogen	0.07

Radiation with energy less than 1 MeV/nm can not penetrate a space suit of 1 mm thickness. Al shielding reduces the low boundary to 40 MeV. When a high-energy ion strikes an atom in metal shielding it can produce secondary radiation and there are cases where a small amount of shielding is worse than none at all. Bremsstrahlung can be created (X-rays) by electrons as they interact with spacecraft shielding.

Radiation damage of electronic components in space environment was studied by Boscherine *et al.* 2003[44].

The radiation-induced degradation of polymeric spacecraft materials under protective oxide coatings was studied by Lachance *et al.* 2001[181].

Polyethylene (C_nH_n) is a relatively inexpensive, stable, and, with a low atomic number, an effective shielding material that has been certified for use aboard the ISS. Several designs for placement of slabs or walls of polyethylene have been evaluated for radiation exposure reduction- and it is shown that 20% or mor reduction in dose in the crew quarters is achievable (see Shavers *et al.* 2004, [280]).

7.3.4 Radiation Risks of Manned Space Missions

Activity of men in space considerably increased since the last 50 years and this trend will continue. Therefore, a careful analysis of possible risks due to radiation is essential especially for long duration missions such as ISS stays of serval months, mission to Mars (estimated duration 1000 days) and lunar missions.

The aim of space radiation programs is to obtain a 95% confidence level that a three 180-day missions at the ISS can be accomplished without exceeding career radiation risk limits. The uncertainty in risk prediction of a Mars mission currently is too high and should also be brought to that level. For this purpose Galactic Cosmic Rays and Solar Energetic Particles Events are simulated using high energy heavy ion beams at the NASA Space Research Laboratory (NSRL) in Brookhaven National Laboratory (BNL) (Schimmerling and Cucinotta, 2004 [273]). NASA Space Research Laboratory became operational in 2003. It is estimated that for each year that astronauts spend in space, about one-third of their DNA will be hit directly by heavy ions.

Chapter 8

Magnetosphere, Ionosphere, Space Weather

A good introduction to the space environment was given in the books [1, 2].

8.1 General Properties

8.1.1 The Magnetosphere

The Earth's magnetosphere can be defined by the area of space around the Earth that is controlled by the Earth's magnetic field.

To a very crude approximation the Earth's magnetic field is a dipole field. Such fields are well known from bar magnets. How is the magnetic field generated? The internal field is generated by a dynamo process, there is a circulation of liquid metal in the deep Earth and this causes the dipole field (bar magnet) with an inclination of 10^0 to the rotation axis of the Earth. The magnetic poles however do not correspond exactly to the geographic poles and moreover, they are reversed, magnetic north is near the geographic south pole. Note that the geographic location of the poles varies, the magnetic poles wander as much as 15 km every year. The earth's magnetic field strength was measured by Carl Friedrich Gauss in 1835. An exponential decay with a half-life of about 1400 years can be clearly seen in the measurements. From Lava flow studies we can deduce that there have been field reversals in the past- the last reversal has occurred 800 000 ago.

88% of the total field can be deduced from a dipole and the magnetic induction $\mathbf{B_D}$ can be derived from dipole potential Φ_D:

$$\mathbf{B}_D = -\mu_0 \nabla \Phi_D = -\nabla \frac{\mu_0 \mathbf{M}.\mathbf{r}}{4\pi r^3} \tag{8.1}$$

[1] Tascione T F, "Introduction to the Space Environment", Krieger (Florida, 1994)

[2] Bergmann-Schaefer, Lehrbuch d. Experimentalphysik, Erde und Planeten. Chapters written by S.F. Bauer and H.O. Rucker

There are the three components of the magnetic induction

$$B_{D,r} = -\frac{\mu_0 M}{4\pi}\frac{2\cos\theta}{r^3} \tag{8.2}$$

$$B_{D,\theta} = -\frac{\mu_0 M}{4\pi}\frac{\sin\theta}{r^3} \tag{8.3}$$

$$B_{D,\Phi} = 0 \tag{8.4}$$

Here, M denotes the magnetic moment of the dipole, for the Earth $M_E \sim 8 \times 10^{22}\,\mathrm{Am}^{-2}$. θ is the angle between the dipole moment and the radiusvector $\mathbf{r}$.

From the above equations we can derive the magnitude of the magnetic induction at $\mathbf{r}$:

$$B(r,\theta) = \sqrt{B_r^2 + B_\theta^2} \tag{8.5}$$

$$B(r,\theta) = \frac{\mu_o M}{4\pi r^3}\sqrt{1+3\cos^2\theta} \tag{8.6}$$

Note that the dipole strength decreases by r^{-3}. In the plane of the equator $\theta = 0$ and:

$$B_{\mathrm{eq}} = \frac{\mu_0 M}{4\pi r^3} = B_0\left(\frac{r_E}{r}\right)^3 \tag{8.7}$$

The induction on the geomagnetic equator is $B_0 \sim 31000\,\mathrm{nT}$ and $r_E = 6378\,\mathrm{km}$ the radius of the Earth.

The space enclosed by the magnetosphere is not empty but filled with trapped particles, namely ions and electrons. The magnetic forces are much stronger than gravity.

The real shape of the boundary of the magnetosphere, the *magnetopause*, is strongly modified by the solar wind. The distance of the magnetopause is

- on the side facing the Sun: 10-12 r_E [3].
- over the poles: 15 r_E
- on the night side the tail reaches past several 100 r_E.

There exists also a neutral gas envelope of the Earth, the Geocorona that extends to 4-5 r_E.

At the side facing the Sun there must be an equilibrium between two pressures:

- pressure of solar wind, this depends on the number of particles n, α which is the angle between velocity of the particles and the normal to the magnetopause and f which denotes the momentum transfer factor. The number of particles per time and area is $nv\cos\alpha$ and the change of momentum $fmv\cos\alpha$. $f = 2$ means total reflection.
- the magnetic pressure at the magnetopause is $B^2/2\mu_0$.

[3] all distances are given in units of the Earth's radius and measured from the Earth's center

Thus the equilibrium conditions becomes:

$$fnmv^2 \cos^2 \alpha = \frac{B^2}{2\mu_0} \tag{8.8}$$

A further condition is that the normal component of the magnetic field is zero:

$$B_n = 0 \tag{8.9}$$

From 8.8 and the typical values for the solar wind particles: $n = 10^{-7}\,\mathrm{m}^{-3}$, $v = v_{\mathrm{SW}} = 300\,\mathrm{kms}^{-1}$, and $r = 10 r_E$ (subsolar point),we get:

$$r = r_E \left(\frac{B^2}{2\mu_0 fnmv_{rmSW}^2} \right)^{1/6} \tag{8.10}$$

And we find that $B \sim 87\,\mathrm{nT}$ are required in order to produce the magnetic pressure needed. If we compute the field from 8.7 the value is $B_D = 31\,\mathrm{nT}$. Thus an amplification is needed. This amplification of the Earth's field is provided by the Chapman-Ferraro currents. Incoming charged particles cannot penetrate the field lines and are deflected which causes currents and these currents produce magnetic fields. Maxwell has shown that when a perfectly conducting flat plane approached a dipole, its externally induced field was the same as the field of an equal image dipole.

A simple drawing illustrates the main components of the magnetosphere (Figs. 8.1, 8.2 and 8.3)

The different parts of a magnetosphere are:

1. *bow shock*: in this front region solar wind particles hit the magnetosphere. The solar wind particles have Mach numbers > 1 that means they are supersonic. This is valid for both the Alfvén and the sound velocity:

$$v_A = \frac{B}{\sqrt{\mu_0 \rho}} \qquad v_S = \sqrt{\frac{c_p}{c_v} \frac{p}{\rho}} \tag{8.11}$$

   and the corresponding Mach numbers are $M_S = v_{\mathrm{SW}}/v_s$; $M_A = v_{\mathrm{SW}}/v_A \sim 10$.

2. The region between the bow shock and the magnetopause is called *magnetosheath*. Here the particles become thermalized- kinetic energy is converted to thermal energy and the plasma is highly turbulent there.

3. The solar wind stretches the dipole field, compressing it on the side towards the sun and stretching it into a long *tail* region. The field lines close at very large distances ($\sim 3000\,R_E$).

4. *plasmasheet*: this is a sheet of plasma in the tail region dividing the two lobes of the Earth's magnetic field. For both electrons and protons the particle density is 0.5 cm^{-3}.

5. *lobes*: they are in the magnetotail have opposite direction and are separated by the plasmasheet -otherwise they would cancel.

6. *plasmasphere*: a torus shaped region, surrounding the Earth. It was detected in 1963 and has a very sharp edge at the *plasmapause* extending to 4-6 Earth radii. It can be also regarded as an extension of the ionosphere. Inside the plasmapause geomagnetic field lines rotate with the Earth. Outside the plasmasphere, magnetic field lines are unable to corotate, the solar wind influence is too large. The plasmasphere is mainly composed of hydrogen.

7. *Van Allen radiation belts*: in 1958 Van Allen discovered the radiation belts; like the plasmasphere they are toroidally shaped. The inner radiation belt extends from 400 to 12000 km above the Earth, the outer belt from 12000 to 60000 km.

In order to understand the dynamics of the current system, we recapitulate the motions of charged particles in a magnetic field:

1. spiral motion: circling about magnetic field lines; Charged particles cannot easily move across magnetic field lines but are forced to spiral around them. Electrons encircle the field line in one direction, ions in the other direction.

2. Bounce motion: the particles move along the field lines from pole to pole. Near the poles they become reflected (since the magnetic field line density is large).

3. drift motion: Curvature of the magnetic field lines and the non-uniform strength of the magnetic field force particles to drift around the earth, ions in one direction, electrons in the other. For the Earth as seen from Europe: Ions go west, electrons east.

In a magnetic field particles are being transported and this causes currents. Due to the currents magnetic fields are generated. In a magnetosphere there are three distinct current systems:

1. Chapman Ferraro currents: they enclose and confine the magnetosphere and are found in the vicinity of the magnetopause.

2. cross tail currents: pass through the center of the magnetotail causing the current sheet.

3. Field aligned currents: transient currents, short circuit through a planet's ionosphere and cause aurorae.

How is the magnetosphere influenced by the solar wind?

- The interaction of enhanced solar wind pressure on the dayside cause a strong reduction of the magnetopause even below the geostationary orbit (6.6 r_E). The observed variations of the distance of the dayside magnetopause are in the range 4.5 to 20 r_E.

- The magnetic moment of the interplanetary magnetic field (magnitude and orientation) determines the size and extension of the magnetosphere.

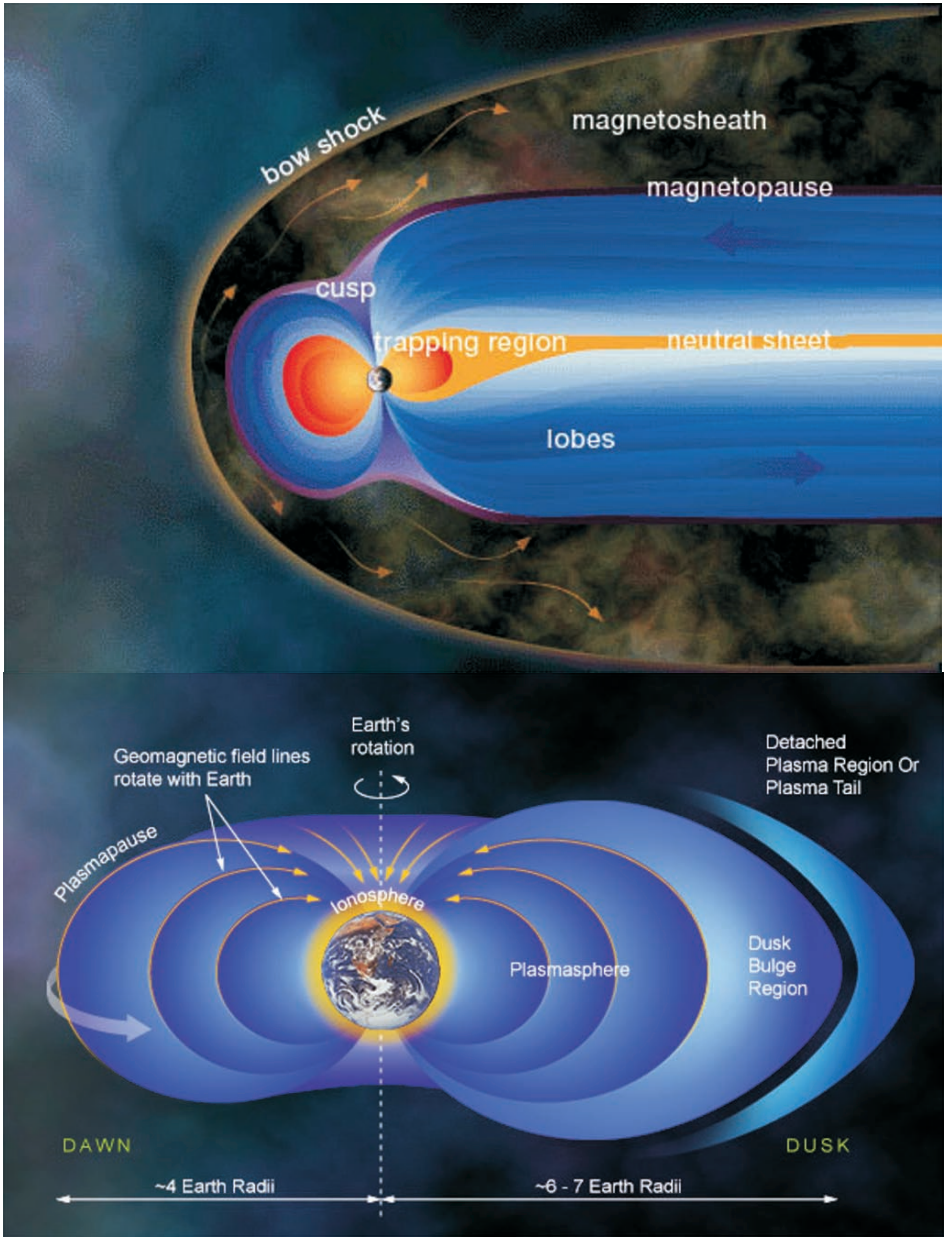


Figure 8.1: The Earth's magnetosphere (above) and the plasmasphere (below). NASA

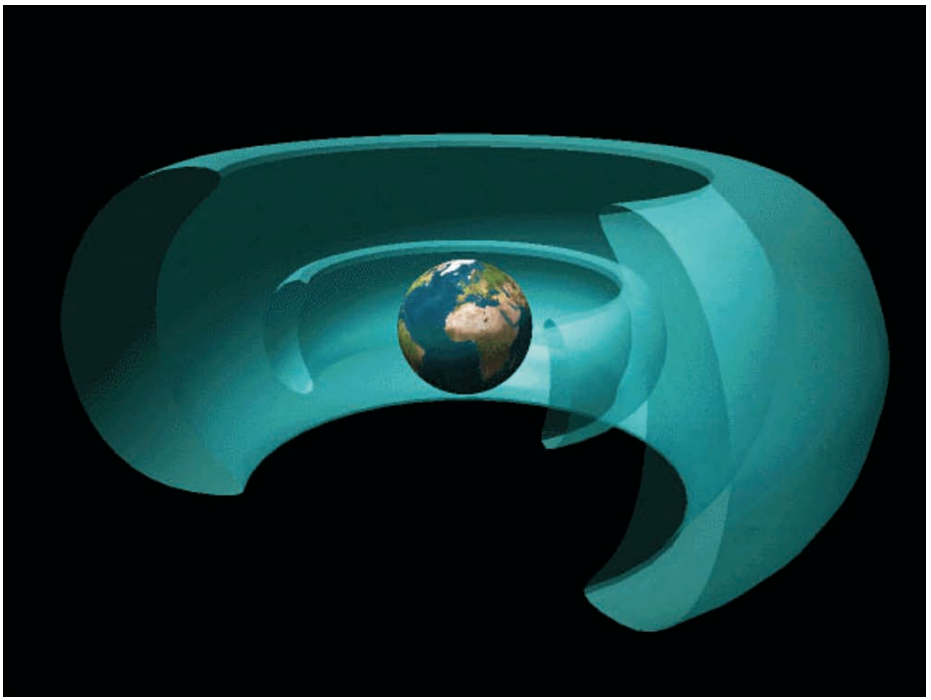


Figure 8.2: The inner and outer Van Allen radiation belt. NASA

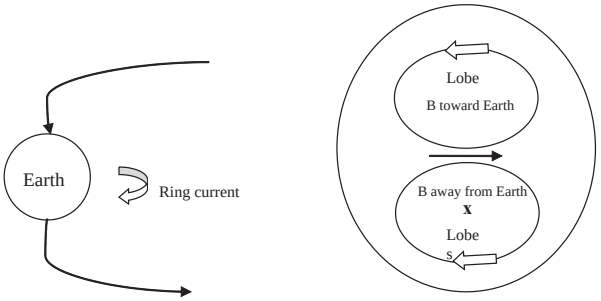


Figure 8.3: The global structure of the Earth's magnetic field. In the two lobes the field is opposite and the lobes are separated by a plasmasheet. For the southern half of the magnetosphere the current is clockwise, for the northern part it is counterclockwise. In the middle both systems add to form a neutral sheet. The right drawing is a cross section of the left at a distance of 20 r_E.

8.1.2 The Ionosphere

The ionosphere contains only a small fraction of the Earth's atmosphere (less than 1 % of the mass of the atmosphere). However, this layer is extremely important for modern telecommunication systems since it influences the passage of radio waves. There are many textbooks describing the complex wave processes, chemical processes, energy deposition and transfer rates in this layer. [4] Because of its name we can expect that the atoms are ionized there. On the sunlit side of the Earth the shorter wavelengths of solar radiation (extreme UV and X rays) are energetic enough to produce ionization of the atoms. Therefore, this layer becomes an electrical conductor supporting electric currents and radio wave propagation.

Historically, it has been divided into regions with specific ionizations.

- lowest D region: between 50 and 90 km.
- E region: between 90 and 150 km,
- F region: contains the F1 and F2 layers.

Ionograms are recorded tracings of reflected high frequency radio pulses generated by an ionosonde. There exist relationships between the sounding frequency and the ionization densities which can reflect it. As the sounder sweeps from lower to higher frequencies, the signal rises above the noise of commercial radio sources and records the return signal reflected from the different layers of the ionosphere.

The top of the ionosphere is at about 1000 km, however there exists no definite boundary between plasma in the ionosphere and the outer reaches of the Earth's magnetic field. In the E region the most important ions are O_2^+, NO^+, in the F region it is O^+. In the F2 layer (at about 400 km) the electron concentration reaches its highest values which is important for the telecommunication systems. At high latitudes there is another source of ionization of the ionosphere– the aurora (see next chapter).

The so called transition height starts at the height of the maximum density of the F2 layer of the Ionosphere and extends upward with decreasing density to a transition height where O^+ ions become less numerous than H^+ and He^+. The transition height depends on day and night:

- daytime: $\sim$ 800 km
- nighttime $\sim$ 500 km.

Above the transition height, the weak ionization has little influence on radio signals.

Some ionospheric parameters are listed in Table 8.1 where the values N_e and T_e denote electron density and electron temperature. For comparison, the values of the solar corona are also given.

[4] e.g. see Ionospheres : Physics, Plasma Physics, and Chemistry by Robert W. Schunk, Andrew F. Nagy, Alexander J. Dessler (Editor), John T. Houghton (Editor), Michael J. Rycroft (Editor), Cambridge Univ. Press, 2004

Table 8.1: Some parameters of the ionosphere.

Layer	N_e cm^{-3}	T$_c$ K	H (Gauss)
Ionosphere			
D	10^3	200	$\sim 3 \times 10^{-1}$
E	10^5 day, 10^3 night	2-3$\times 10^2$	$\sim 3 \times 10^{-1}$
F_1	10^5 day, absent night	1000	$\sim 3 \times 10^{-1}$
F_2	10^6 day, 10^5 night	1-3$\times 10^3$	$\sim 3 \times 10^{-1}$
Solar Corona	10^4...10^8	$\sim 10^6$	10^{-5}...1

At low latitudes the largest electron densities are found in peaks on either side of the magnetic equator, which is called the equatorial anomaly. Normally one would expect that the peak concentration will occur at the equator because of the maximum in solar ionizing radiation. This peculiarity can be explained by the special geometry of the magnetic field and the presence of electric fields. The electric fields transport plasma and are caused by a polarizing effect of thermospheric winds.

The ionosphere varies because of two reasons:

- two varying sources of ionization (aurora, Sun)
- changes in the neutral part of the thermosphere, which responds to solar EUV radiation.

Thus the ionospheric variation mainly occurs at a 24 h period (daytime-nighttime) and over the 11 year cycle of solar activity. We observe considerable changes in the F-region maximum density ($N_{\max}$) of the electrons which influences the plasma frequency that is proportional to it. On shorter time scales solar X-ray radiation changes dramatically during a solar flare eruption. This effect increases the D and E ionization. During a geomagnetic storm the auroral source of ionization becomes more intense. In extreme cases aurorae can be seen at moderate latitudes (Italy, Mexico). Another source of variability in the ionosphere comes from the interaction of charged particles with the neutral atmosphere in the thermosphere. Thermospheric winds can push the ionosphere along the inclined magnetic field line to a different altitude. Moreover the composition of the thermosphere affects the rate that ions and electrons recombine. During a geomagnetic storm energy input at high latitudes produces waves and changes in the thermospheric winds and composition. The electron concentration can increase (positive phases) and decrease (negative phases). The ionospheric variability is given in Table 8.2.

HF communication depends on radio waves that are reflected in the ionosphere. This is characterized by the maximum usable frequency (MUF) and the lowest usable frequency (LUF). The MUF depends on the peak electron density in the F region and the angle of incidence of the emitted radio wave. As we have seen, this changes during the day, over the solar cycle and during geomagnetic disturbances. The LUF is controlled by the amount of absorption of the radiowave in the lower D and E layers. This is severely affected by solar flares. All single frequency GPS

Table 8.2: Variation of the ionosphere

Ionospheric parameter	Variation Diurnal (Mid-Latitude)	Variation Solar Cycle (daytime)
$N_{\max}$	$1 \times 10^5 ... 1 \times 10^6 e^-/\mathrm{cm}^3$	$4 \times 10^6 .... 2 \times 10^6 e^-/\mathrm{cm}^3$
Max. Usable Freq. MUF	Factor of 10 12...46 MHz Factor of 3	Factor of 5 21 ...42 MHz Factor of 2
Total Electron Content TEC	$5...50 \times 10^{16} e^-/\mathrm{m}^2$ Factor of 10	$10...50 \times 10^{16} e^-/\mathrm{m}^2$ Factor of 5

receivers must correct the delay of the GPS signal as it propagates through the ionosphere to the GPS satellite (at 22 000km altitude).

The maximum usable frequency depends on the angle of the wave relative to the horizon.

The ionosphere may become highly turbulent, mainly in the high latitude and low latitude F region and at special times (often after sunset). In this context turbulence is defined as small scaled structures (scale length cm to m) which are irregular and embedded in the large scale ambient ionosphere (tens of kilometers). In the equatorial region plasma irregularities are generated just after sunset and may last for several hours. At high latitudes these irregularities may be generated during day and night. Both effects occur most frequently during the solar cycle maximum. Radio signals become disrupted by these perturbations and the effect is known as *ionospheric scintillation*. The bigger the amplitude of the scintillated signal the greater the impact on communication and navigation systems.

8.2 Solar Activity and Magnetosphere

C.F. Gauss (1777-1855) measured variations of the terrestrial magnetic field. By the end of the 19th century it was recognized that some disturbances of the Earth's magnetic field could be traced to the Sun. Some were found to be related to solar flares, others showed a 27 day recurrence interval which also points to a solar origin. In 1930 Sidney Chapman and Vincent Ferraro proposed that the Sun sent out huge clouds of electrically neutral plasma, and that magnetic storms arose when those clouds enveloped the Earth. The strong field of the Earth would hold off the cloud, carving a cavity in the cloud in which the Earth and its magnetic field would be confined (see drawing from their 1931 article Fig. 8.5). They also speculated that a ring current would then be set up, though they had no clear idea of the way it happened.

Today we know that the flow of plasma from the Sun is not confined to isolated clouds, but goes on all the time, in the form of the solar wind. Denser and faster clouds arising from coronal mass ejections (see corona) were later identified as the real cause of sudden commencements.

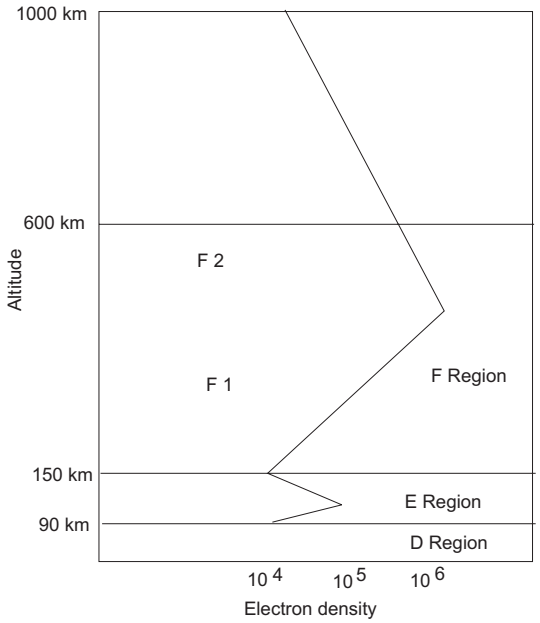


Figure 8.4: The Earth's ionosphere

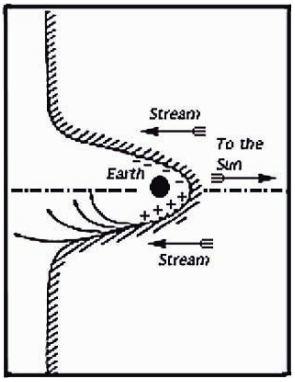


Figure 8.5: Original drawing of Chapman and Ferraro showing the interaction of plasma from the Sun and the Earth's magnetic field.

Generally the solar wind arriving at the Earth's magnetopause has the following pressure components:

$$\text{Dynamic} \rightarrow \rho v^2 \tag{8.12}$$

$$\text{Static} \rightarrow nkT \tag{8.13}$$

$$\text{Magnetic} \rightarrow B^2/2\mu_0 \tag{8.14}$$

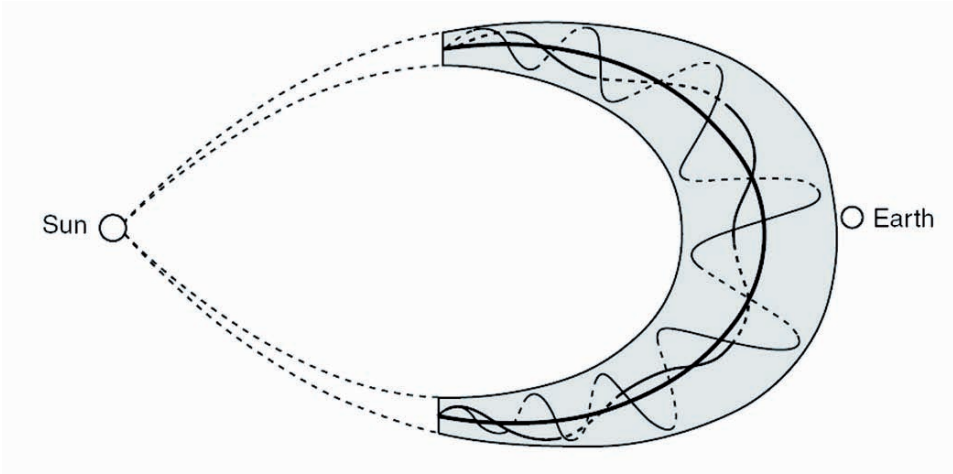


Figure 8.6: An interplanetary coronal mass ejections interacts with the Earth magnetic field (from http://www-ssc.igpp.ucla.edu/gem/tutorial/2000Russell.pdf)

The pressure applied by the solar wind to the magnetopause varies with the angle of the normal to the solar wind flow. The pressure is dominated by the dynamic pressure. At the magnetopause the dynamic pressure is zero and the static pressure dominates. Inside the magnetosphere the pressure is dominated by the magnetic pressure.

The so called standoff distance i.e. the distance of the magnetopause is given by:

$$L_{\rm mp} = 107.4(n_{\rm SW} v_{\rm SW}^2)^{-1/6} \tag{8.15}$$

The interaction between the interplanetary magnetic field and the Earth's magnetic field depends on the orientation of the former with respect to the Earth's field. This was studied first by Dungey (1961 [82]) and is called Dungey's model.

The pressure of the solar wind rises and falls. The reacting of the magnetopause is a shrinking or expansion. When the boundary is hit by a fast flow from a CME, the shrinking can go beyond the geosynchronous orbit of satellites (at 6.6 R_E). As it is seen from the drawing (Fig. 8.7), a southward oriented IMF is recognized as the most important factor promoting storms and substorms in the magnetosphere (Fairfield and Cahill, 1966 [95]).

When the interplanetary magnetic field is oriented southward, then a flow of plasma is predicted to the dayside of the magnetosphere after reconnection in the tail.

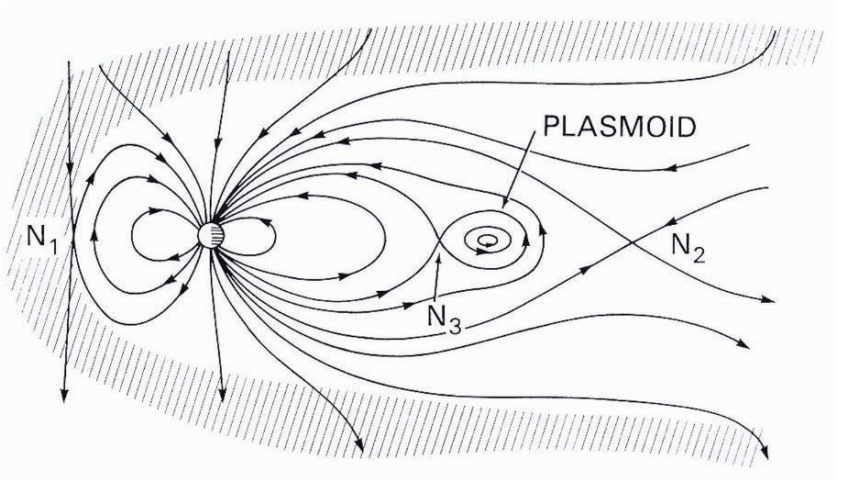


Figure 8.7: Interaction of a southward oriented IMF with the magnetosphere. Possible reconnection points are denoted by N_1, N_2, N_3. Also the formation of a disconnected plasmoid is indicated. From "A Brief History of Magnetospheric Physics during the Space Age" by D.P. Stern

8.2.1 Magnetic Storms

The Sun heats the Earth's atmosphere. Also the degree of ionization in the ionosphere increases at the dayside and this causes convection in the ionosphere. By this convection charged particles are transported into the magnetosphere and by dynamo action ionospheric electric currents above the equator up to mid latitudes are generated. These currents produce a magnetic field which moves with the subsolar point. So there is a 12 h variation for a given observing site in the measurements of the field strength.

The Sun emits particles and the solar wind compresses the magnetosphere as it has been mentioned before. High speed particles further compress the magnetosphere, and a magnetic storm begins with a SSC (storm sudden commencement). The number of charged particles trapped within regions of the magnetosphere (radiation belts) is increased. These particles drift around the Earth creating a ring current that produces a depression of the horizontal magnetic field, seen at lower latitudes around the world as a magnetic storm. This is followed by the recovery phase, lasting one day or more, during which the ring current subsides and the magnetic field returns to normal. Charged particles are guided down the field lines into the upper atmosphere. This produces auroral electrojets (large horizontal currents that flow in the D and E regions of the auroral ionosphere) which are intense east-west currents. Associated with these currents are intense magnetic fields causing magnetic disturbances observed there.

8.2.2 Particles and Particle Motion

The solar wind sweeps toward Earth at super sonic speeds ranging from 300 to 1000 km/s. It distorts the Earth's magnetic field which forms out a comet shaped magnetosphere.

There are two Van Allen belts of particle concentration : a) small inner belt between 1 and 2 Earth radii where protons of energy 50 MeV (see also Table 8.3) and electrons with energies > 30 MeV reside and b) outer larger belt from 3 to 4 Earth radii where less energetic protons and electrons are concentrated. The inner belt is relatively stable, the outer belt varies in its number of particles by as much as a factor of 100.

Charged particles trapped in the belts spiral along the field lines while bouncing between the northern and southern mirror points. Particles in the inner belt may interact with the upper atmosphere causing the auroral oval which is an annulus centered over the magnetic poles and around 3000 km in diameter during quiet times. The location of the auroral oval is usually found between 60 and 70 degrees of magnetic latitude (north and south).

When charged particles follow magnetic field lines a current flows, this is called a Birkeland current [5]. Today, often the term auroral electrojets is used. Auroral Birkeland currents can reach about 10^6 A and heat up the upper atmosphere which results in increased drag on low-altitude satellites.

Table 8.3: Typical Particle Energies

0.03 eV	Molecule of oxygen or nitrogen in the air
0.5	Atom or molecule $T_\odot$, surface
0.67 eV	Proton or neutron escape the Earth's gravity
1000 - 15,000 eV	Electron in the polar aurora
1.4 MeV	Energy of electrons from radioactive potassium major source of the Earth's heat
10-100 MeV	Proton energies in the inner radiation belt
10-15,000 MeV	Range of energies in solar flares
1-100,000,000,000 GeV	Cosmic ray ions; as their energy goes up, their intensity goes down

The interaction of plasma of the solar wind with the Earth's magnetosphere causes currents as shown in Fig. 8.8:

- flow eastwards down the morning side around the polar regions
- flow spacewards in the evening side.

Interplanetary field lines are swept back around the Earth's magnetic field by the solar wind. There is an electric field according to

$$\mathbf{E} = \mathbf{v}_{\mathrm{SW}} \times \mathbf{B}_{\mathrm{SW}}/c \tag{8.16}$$

[5] Birkeland, 1903

This equation follows from Ohm's law ($\mathbf{j} \sim \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ and in the case of a large conductivity $\sigma >> 1$ the term $\mathbf{j}/\sigma \to 0$. This field is from dawn to dusk, there is a field aligned current and particles move from dawn to dusk. Because this circulation is analogous to thermal convection cells, this phenomenon is also called convection electric field.

Finally, one also has to take into account the corotational electric field. This is cause by the rotation of a planet's magnetic field which induces an electric field in the radial direction. The magnetic field moves at

$$\mathbf{v} = \omega_{\text{rot}} \times \mathbf{r} \tag{8.17}$$

and an electric field is induced by Ohm's law (again we consider large conductivity):

$$\mathbf{E}_{\text{cor}} = -\frac{\mathbf{v} \times \mathbf{B}}{c} = -\frac{(\omega_{\text{rot}} \times \mathbf{r}) \times \mathbf{B}_0}{cr^3} \tag{8.18}$$

The motion of charged particles in the magnetosphere is thus caused by a

- drift: due to gradient in the field strength, field curvature
- acceleration due to electric fields along the field lines, field aligned currents

Where do the particles come from? Interplanetary particles can enter the magnetosphere via different processes:

1. spiral down into the polar cusp- there are open magnetospheric field lines there. Atmospheric ionization is enhanced there during enhanced solar activity → aurora.
2. reconnection is an important process. It occurs when the interplanetary field has a component antiparallel to the planetary field. Reconnection leads to neutral points and solar wind particles can enter there. The locations of reconnection are the day side magnetopause and the magnetotail.
3. Kelvin-Helmholtz instability . The fast solar wind flows past the magnteosphere inducing ripples in the magnetospheric boundary. These ripples induce a filed perpendicular to the solar wind flow and thus particles diffuse into the planetary magnetosphere.

Particles are lost because of different processes, such as losses due to the mirror points. For particles with a certain pitch angle, the mirror points lie within the atmosphere and the particle gets lost. Another process is charge exchange of magnetospheric ions.

8.2.3 Aurora

There are many shapes and features of aurorae. They generally start at 100 km above the surface and extend upward along the magnetic field for hundreds of km. Auroral arcs can nearly stand still and then suddenly move (dancing, turning). After midnight one often sees a patchy appearance of aurorae, and the

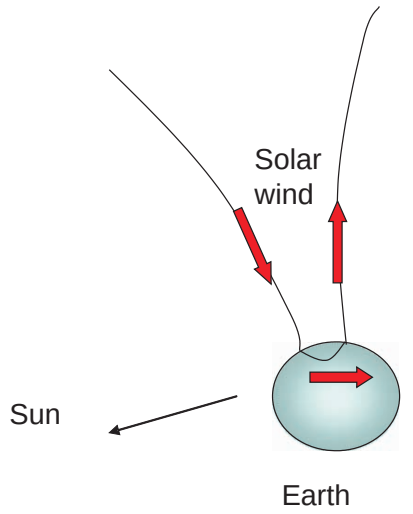


Figure 8.8: Birkeland currents. The currents flow downwards on the morning side and spacewards on the evening side.

patches blink on and off every 10 s or so. Most of aurorae are greenish yellow and sometimes the tall rays turn red at their top and along their lower edge. On rare occasions sunlight hits on the top creating a faint blue color.

The different colors depend on the specific atmospheric gas, its electrical state and on the energy of the particle that hits the atmospheric gas. Atomic oxygen is responsible for the two main colors of green (557.7 nm, at a height below 400 km) and red (630.0 nm, about 400 km or higher). Excited nitrogen also emits light (600-700 nm; below 200 km). Auroral displays are intensified if the interplanetary magnetic field is in the opposite direction to the Earth's magnetic field. The geomagnetic storms produce brightness changes and motion in the aurorae and these are called auroral substorms. Recent models of aurorae explain the phenomenon by a process of release of energy from the magnetotail, called magnetic reconnection. Regions of opposite magnetic fields come together and the magnetic field lines can break and reconnect in new combinations. The point of reconnection in the magnetotail lies usually at 100 Earth radii (see 8.7). When the solar wind adds sufficient magnetic energy to the magnetosphere, the field lines there overstretch and a new reconnection takes place at 15 Earth radii, the field collapses and electrons are injected into the atmosphere.

Reconnection stores large amounts of energy in the Earth's magnetic field until it is released explosively. The cycle of energy storage and release is called substorm. Multiple substorms lead to magnetic storms and acceleration of particles to very high energies. These particles damage satellites.

The geomagnetic field is measured by magnetometers and the data are often given as 3-hourly indices that yield a quantitative measure of the level of geomagnetic activity. The K-index is given from 0 to 9 and depends on the observing

Table 8.4: Corrected magnetic latitudes of some cities

Atlanta	44.5	Athens	31.3	Adelaide	45.9
Boston	51.7	Berlin	48.3	Buenos Aires	23.3
Chicago	52.2	Copenhagen	51.9	Capetown	41.5
Dallas	42.7	Edinburgh	53.0	Christchurch	49.9
Denver	48.3	London	47.5	Comodoro Rivadavia	32.1
Great Falls, MT	54.9	Madrid	33.3	Concepcion, Chile	23.2
Havana	34.1	Moscow	51.8	Dunedin	53.0
Los Angeles	39.8	Paris	44.2	Durban	38.8
Mexico City	29.1	Perm	53.8	East London	41.1
Minneapolis	55.1	Prague	45.5	Hobart	53.6
New York	50.6	Rome	35.5	Melbourne	48.4
Quebec City	56.2	St. Petersburg	56.1	Perth	43.9
San Francisco	42.5	Warsaw	46.7	Punta Arenas, Chile	38.6
Seattle	52.7	Beijing	34.1	Sydney	43.5
St. Louis	49.2	Irkutsk	47.0	Toronto	53.9
Seoul	31.0	Washington, DC	49.1	Tokyo	29.0
Winnipeg	59.5	Vladivostok	36.5	Vienna	43.0

Table 8.5: Extension of the auroral zone. The first values given is the magnetic latitude (Lat), the second the Kp index.

Lat	Kp	Lat	Kp	Lat	Kp	Lat	Kp	Lat	Kp
66.5	0	64.5	1	62.4	2	60.4	3	58.3	4
56.3	5	54.2	6	42.2	7	50.1	8	48.1	9

station. The globally averaged K_p index is a measure for the global auroral activity.

When geomagnetic activity is low, the aurora typically is located at about 67 degrees magnetic latitude, in the hours around midnight. As activity increases, the region of aurora expands towards the equator. When geomagnetic activity is very high, the aurora may be seen at mid and low latitude locations (see Table 8.4) around the earth that would otherwise rarely experience the polar lights. In Table 8.5 auroral boundaries are given as a function of the Kp index.

The magnetic activity produced by enhanced ionospheric currents flowing below and within the auroral oval is measured by the Auroral Electrojet Index AE. The definition of this index is as follows: at a certain time the total range of deviation from quiet day values of the horizontal magnetic field (h) around the auroral oval. Defined and developed by Davis and Sugiura in 1966, AE has been usefully employed both qualitatively and quantitatively as a correlative index in studies of substorm morphology, the behavior of communication satellites, radio propagation, radio scintillation, and the coupling between the interplanetary mag-

netic field and the Earth's magnetosphere. For these varied topics, *AE* possesses advantages over other geomagnetic indices or at least shares their advantageous properties.

8.2.4 Geomagnetic Indices

Daily regular magnetic field variations arise from current systems caused by regular solar radiation changes. Other irregular current systems produce magnetic field changes caused by

1. the interaction of the solar wind with the magnetosphere,
2. by the magnetosphere itself,
3. by the interactions between the magnetosphere and ionosphere,
4. and by the ionosphere itself.

Therefore, magnetic activity indices were designed to describe variation in the geomagnetic field caused by these irregular current systems.

Let us give a brief description of other geomagnetic indices which are interesting for the solar-terrestrial relations.

DST Index

DST stands for Disturbance Storm Time. The DST is an index of magnetic activity derived from a network of near-equatorial geomagnetic observatories that measures the intensity of the globally symmetrical equatorial electrojet (the "ring current"). Thus DST monitors the variations of the globally symmetrical ring current, which encircles the Earth close to the magnetic equator in the Van Allen (or radiation) belt of the magnetosphere. During large magnetic storms the signature of the ring current can be seen in ground magnetic field recordings worldwide as so-called main phase depression. The ring current energization which results in typical depression of 100 nT is related to magnetic reconnection processes at the neutral sheet.

Kp, Ap and C Index

The *K-Index* was first introduced by J. Bartels in 1938. It is a quasi-logarithmic local index of the 3-hourly range in magnetic activity relative to an assumed quiet-day curve for a single geomagnetic observatory site. The values consist of a single-digit 0...9 for each 3-hour interval of the universal time day (UT).

The planetary 3-hour-range index *Kp* is the mean standardized K-index from 13 geomagnetic observatories between 44 degrees and 60 degrees northern or southern geomagnetic latitude. The scale is 0...9 expressed in thirds of a unit, e.g. 5- is 4 2/3, 5 is 5 and 5+ is 5 1/3. This planetary index is designed to measure solar particle radiation by its magnetic effects. The 3-hourly Ap (equivalent range) index is derived from the Kp index (see Table 8.6). This table is made in such a way that at a station at about magnetic latitude 50 degrees, Ap may be regarded

Table 8.6: Transformation between the Kp and the Ap index

Kp =	0o	0+	1-	1o	1+	2-	2o	2+	3-	3o	3+	4-	4o	4+
Ap =	0	2	3	4	5	6	7	9	12	15	18	22	27	32
Kp =	5-	5o	5+	6-	6o	6+	7-	7o	7+	8-	8o	8+	9-	9o
Ap =	39	48	56	67	80	94	111	132	154	179	207	236	300	400

Table 8.7: Transformation between the Ap and the Cp index

Cp	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Ap	2	4	5	6	8	9	11	12	14	16	19
Cp	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	
Ap	22	26	31	37	44	52	63	80	110	160	

as the range of the most disturbed of the three field components, expressed in the unit of 2 g. A daily index *Ap* is obtained by averaging the eight values of Ap for each day. The *Cp* index, the daily planetary character figure, is defined on the basis of Ap according to Table 8.7

Another index devised to express geomagnetic activity on the basis of the Cp index is the *C*9 index which has the range between 0 and 9. The conversion table from the Cp index to the C9 index is given by 8.8

AE and Other Indices

These indices describe the disturbance level recorded by auroral zone magnetometers.

In order to determine these indices, horizontal magnetic component recordings from a set of globe-encircling stations are plotted to the same time and amplitude scales relative to their quiet-time levels. They are then graphically superposed. The upper and lower envelopes of this superposition define the AU (amplitude upper), the AL (amplitude lower) indices and the difference between the two envelopes determine the *AE* (Auroral Electrojet) index, i.e., AE = AU - AL. AO is defined as the average value of AU and AL.

A summary of the indices as well as a few other indices can be found in Table 8.9

Table 8.8: Transformation between the Cp and the C9 index

Cp	0.0-0.1	0.2-0.3	0.4-0.5	0.6-0.7	0.8-0.9
C9	0	1	2	3	4
Cp	1.0-1.1	1.2-1.4	1.5-1.8	1.9	2.0-2.5
C9	5	6	7	8	9

Table 8.9: Summary of geomagnetic indices

aa	3-hour range index, derived from two antipodal stations
AE, AU, AL	1- , 2.5-minute, or hourly auroral electrojet indices
am, an, as	3-hour range (mondial, northern, southern) indices
Ap	3-hour range planetary index derived from Kp
C, Ci, C9	Daily local (C) or international (Ci) magnetic character; C9 was first derived from Ci, then from Cp
Cp	Daily magnetic character derived from Kp
Dst	Hourly index mainly related to the ring current
K	3-hour local quasi-logarithmic index
Km	3-hour mean index derived from an average of K indices (not to be confused with the Km of the next item)
Km, Kn, Ks	3-hour quasi-logarithmic (mondial, northern, southern) indices derived from am, an, as
Kp, Ks	3-hour quasi-logarithmic planetary index and the intermediate standardized indices from which Kp is derived (not to be confused with the Ks of the preceding item)
Kw, Kr	3-hour quasi-logarithmic worldwide index and the intermediate from which Kw is derived
Q	Quarter hourly index
R	1-hour range index
RX, RY, RZ	Daily ranges in the field components
sn, ss	3-hour indices associated with an and as
U, u	Daily and monthly indices mainly related to the ring current
W	Monthly wave radiation index

8.2.5 Solar Indices

10.7 cm Radio Flux

The sun emits radio energy with slowly varying intensity. This radio flux, which originates from atmospheric layers high in the sun's chromosphere and low in its corona, changes gradually from day to day in response to the number of spot groups on the disk. Solar flux from the entire solar disk at a frequency of 2800 MHz has been recorded routinely by a radio telescope near Ottawa since February 1947. The observed values have to be adjusted for the changing Sun-Earth distance and for uncertainties in antenna gain (absolute values). Fluxes are given in units of $10^{-22}\,\mathrm{Js^{-1}m^{-2}Hz^{-1}}$.

Sunspot Numbers

The sunspot number index is also often called Wolf number in reference to the Swiss astronomer J. R. Wolf who introduced this index in 1848; details about how to obtain that number can be found in the chapter about sunspots and the solar cycle.

Table 8.10: Navigation systems

System	Frequency	
Omega	VLF, kHz	about 10^4 Hz
Loran-C	LHF	about 10^5 Hz
GPS	UHF, GHz	about 10^9 Hz

8.2.6 Navigation Systems

Modern travel requires exact latitude, longitude and altitude information in real time. Therefore terrestrial based radio wave systems such as the Loran-C and the Omega-system were developed. They use large transmitter antennas to send low-frequency (LF) and very-low-frequency (VLF) radio signals along the ground and off the reflective layer provided by the ionosphere. Thus, vast distances over land and sea can be reached. More recently, space-based systems have become the tools for navigation, among others the GPS system (Global Positioning System). The advantage of space-based systems is that the satellites can easily cover the globe. A user can obtain an accurate three dimensional position (his location and altitude) as soon as at least four satellites are in view.

However both navigational systems, space-systems as well as systems on the surface suffer from the transmission through the ionosphere. The Omega system requires it, the Loran system tries to avoid it and the GPS system depends on radio signals that pass through it. Flares produce X rays and we have already discussed the influence of this shortwave radiation on the D and E region in the ionosphere. Navigation with Loran-C and Omega systems thus is influenced by these events and during the maximum phase of the solar cycle daylight users of Loran-C and Omega systems have more difficulties. The GPS system is not influenced by this perturbation. The GPS operations are affected by the total electron content of the ionosphere along the path to the satellite and are thus influenced by geomagnetic storms. Whereas solar X-rays impact only the sunlit hemisphere of Earth, geomagnetic storms are ubiquitous. The ionospheric response to the storms also depends on the latitude. The conditions nearer to the equator or nearer to the poles vary for the user. It must also be stressed that a quiet undisturbed geomagnetic field does not necessarily dictate an undisturbed equatorial ionosphere. The influence of TEC variation (Total Electron Content) on GPS receivers is smaller for dual band receivers which actually measure the effect of the ionosphere on the GPS signals and correct the resulting positions for these. Unpredictable density enhancements can occur in the evening hours and cause scintillations which affect both dual- and single-frequency GPS receivers.

We summarize the effect of the space environment on the navigation systems:

- Loran-C: Phase and amplitude shifts due to skywave interference at the limits of coverage area.
- GPS: Carrier loss-of-lock due to ionospheric density fluctuations with solar or geomagnetic activity.

- Omega: Phase anomalies due to varying ionospheric reflection height; caused by solar or geomagnetic activity.

Example of a Case Study

We want to give an example of combined observations of the Sun-Earth system. This is extracted from the work of Hanuise *et al.* 2006 [126]. Flares and CMEs (especially halo CMEs) were observed in solar active region AR 10365. On May 27 and 28 three halo CMEs were observed. On May 29 the disturbance propagated to L1 and was measured as two shocks and pulses by the spacecraft ACE. The magnetosphere became strongly compressed and the sub-solar magnetopause moved inside five Earth radii. This causes a geomagnetic storm with several impacts:

- expansion of the auroral oval, and aurorae seen at mid latitudes
- significant modification of the total electron content in the sunlight high-latitude ionosphere,
- perturbation of radio-wave propagation → HF blackouts and increased GPS signal scintillation,
- heating of the thermosphere → increased satellite drag.

8.2.7 Radio Communication

The ionosphere affects the propagation of radio signals in different ways depending on their frequencies. Frequencies below ∼50 MHz are reflected in the ionosphere; this allows radio communication to distances of many thousands of kilometers. Radio signals at frequencies above 50 MHz penetrate the ionosphere and are useful for ground-to-space communications. Frequencies between 2 and 30 MHz are affected by increased absorption, higher frequencies by different reflection properties in the ionosphere (see Fig. 8.9).

Reflection in the ionosphere allows short wave radio reception to occur beyond the limits of line of sight. It is utilized by amateur radio enthusiasts, shortwave broadcast stations (such as BBC and Voice of America) and others and AM stations (Mittelwelle).

Three frequencies are important for the propagation of radiowaves in the ionosphere:

- The limiting frequency at or below which a radio wave is reflected by an ionospheric layer at vertical incidence if given by:

$$f_{\text{crit}} = 9 \times 10^{-3} \sqrt{N_e} \tag{8.19}$$

  N_e is the electron density cm^{-3} and the frequency f_{crit} is in MHz.

- From this frequency we can deduce the maximum usable frequency MUF by:

$$f_{\text{MUF}} = \frac{f_{\text{crit}}}{\sin \alpha} \tag{8.20}$$

  where α is the angle of the wave relative to the horizon.

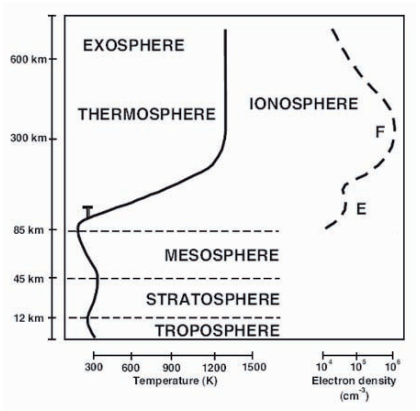


Figure 8.9: Different layers in the ionosphere. Reflection of radio waves occur in the E-layer at 110 km and also the F layers (170 km, 250 km) reflect waves. In the D-layer an absorption occurs. Within the auroral oval the nighttime E layer plasma densities can be much higher.

- The cutoff frequency is the frequency below which a radio waves fails to penetrate the ionosphere.

TV and FM radio stations (on VHF) are affected little by solar activity. HF ground to air, ship to shore, amateur radio etc. are affected strongly. Also the Faraday rotation of the plane of polarization has to be taken into account (for satellite which employ linear polarization up to 1 GHz).

During a solar flare event a sudden increase of X-ray emission causes a large increase in ionization in the lower regions of the ionosphere on the sunlit side of the Earth. Very often one observes a sudden ionospheric disturbance (SID). This affects very low frequencies (OMEGA) as a sudden phase anomaly (SPA) or a sudden enhancement of the signal (SES). At HF and sometimes also at VHF an SID may appear as a short wave fade (SWF). Depending on the magnitude of the solar flare such a disturbance may last from minutes to hours. At VHF the radio noise created by solar flares interferes with the signal. The occurrence of solar flare is modulated by the solar activity.

Flares may also emit energetic particles. The PCA (polar cap absorption) is caused by high energetic particles that ionize the polar ionosphere. A PCA may last from days to weeks depending on the size of the flare and the interaction of the high energetic particles emitted by the flare and the Earth's magnetosphere. During these events polar HF communication becomes impossible. A coronal mass ejection may be a consequence of a large solar flare or a disappearing filament and is an ejection of a large plasma cloud into the interplanetary space. Such a coronal mass ejection (CME) travels through the solar wind and may also reach the Earth. This results in a global disturbance of the Earth's magnetic field and is known as a geomagnetic storm. High speed solar wind streams originating in coronal holes on the Sun's corona hits the Earth's magnetosphere and also causes ionospheric disturbances.

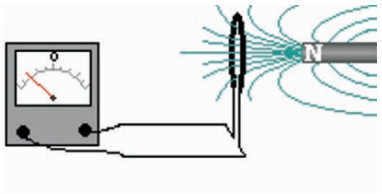


Figure 8.10: Principle of electromagnetic induction. When moving a bar magnet a current is induced and can be measured.

Dudeneye *et al.*,1986 [81] discussed criteria for the development of ionosphere electron concentration vertical profile. The ways in which the ionosphere influences the properties of a radio signal are reviewed in Bradley, 1984 [47].

8.2.8 Geomagnetically Induced Currents

The coupling between the magnetosphere and the ionosphere leads to ionospheric electric fields. At low latitudes the ionospheric plasma is co-rotating with the Earth. At large latitudes convection occurs (Harang discontinuity).

Ground effects of space weather are generally known as GIC (geomagnetically induced currents). A real time GIC simulator is available at

http://www.spaceweather.gc.ca/gic_simulator_e.php.

The changing magnetic field induces currents in the Earth itself- the induced currents produce magnetic fields that again disturb the magnetic field at the Earth's surface. The magnitude of the induced currents and electrical fields depends on electrical conductivities of the different layers within the Earth. Magnetic variations with lower frequencies penetrate deeper.

These currents are driven by the geoelectric field associated with a magnetic disturbance in electric power transmission grids, pipelines, communication cables and railway equipment. GIC are dc currents. They may cause several effects because they increase existing current and this may cause saturation:

- Increase of harmonics,
- unnecessary relay trippings,
- increase in reactive power loss,
- voltage drops,
- permanent damage to transformers,
- black out of the whole system.

When flowing from the pipeline into the soil, GIC may increase corrosion of the pipeline, and the voltages associated with GIC disturb the cathodic protection system and standard control surveys of the pipeline.

On March 13, 1989, the most famous GIC failure occurred in the Canadian Hydro-Quebec system during a great magnetic storm. The system suffered from a nine-hour black-out.

A theoretical calculation of GIC in a given network (power grid, pipeline etc.) can be divided into two steps:

- Calculate the geoelectric field created primarily by ionospheric-magnetospheric currents and affected secondarily by the earth's conductivity distribution. This is also called the geophysical step.
- Calculate the currents produced by the geoelectric field in the circuit system constituted by the network and its earthings.

The first step is generally more difficult, partly because the space and geophysical input parameters are not well known.

The effects of geomagnetic disturbances on electrical systems at the earth's surface were studied e.g. by Boteler *et al.* (1998 [45]). A prediction of Geomagnetically Induced Currents in Power Transmission Systems was given by Pirjola *et al.* (2000 [246]).

A short description of the vulnerable Swedish power system (because being close to the auroral oval) and pipeline system together with a historical description of the effects that occurred at times of geomagnetically induced currents (GICs), up to the Halloween events in 2003 and event in November 2004 was done by Lundstedt, 2006 [203].

On 30 October 2003 50 000 customers in Southern Sweden had no electricity due to a power failure caused by a GIC (see Pulkkinen *et al.*, 2005 [250]). Research on historical geomagnetic storms can help to create a good data base for intense and super-intense magnetic storms. For the event on March 13, 1989 the Dst=-640 nT. Lakhina *et al.* 2005 [184] claimed to have found evidence for a superstorm that occurred on Sep 1-2 1859 with a Dst=-1760.

8.2.9 Systems Affected by Solar or Geomagnetic Activity

In this paragraph we give a summary of the influence of solar and geomagnetic activity (driven by solar events) on various systems.

- HF Communications
  - Increased absorption
  - Depressed MUF
  - Increases LUF
  - Increases fading and flutter
- Surveillance Systems
  - Radar energy scatter (auroral interference)
  - Range errors

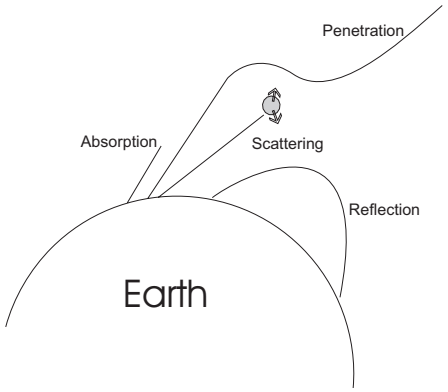


Figure 8.11: Radio signal propagation in the ionosphere.

  - Elevation angle errors
  - Azimuth angle errors

- Satellite Systems
  - Faraday rotation
  - Scintillation
  - Loss of phase lock
  - Radio Frequency Interferences (RFI)

- Navigation Systems
  - Position errors

8.2.10 The Global Ionosphere-Thermosphere Model

In this section we shortly outline the **G**lobal **I**onosphere-**T**hermosphere **M**odel, GITM. The model is described in Ridley, Deng and Toth, 2006 [258]. A three dimensional spherical grid is used that can be stretched both in latitude and altitude. The resolution is fixed in longitude. GITM is flexible and different models of high-latitude electric fields, auroral particle precipitation, solar EUV inputs, and particle energy deposition can be used. The magnetic field can be represented by an ideal dipole magnetic field or a more realistic complex magnetic field. Many of the source terms can be controlled (switched on and off, or values set).

The coupling between the ionosphere-thermosphere is extremely important for space weather applications, such as to study the drag on satellites due to heating of the atmosphere, GPS degradation analysis and examine the interaction of these layers with the lower atmosphere and thus the impact on climate. Concerning the

solar EUV heating the GITM calculates an altitude dependent heating efficiency. The heating efficiency starts with a value of 0.25 at a height of 100 km, reaches a maximum of 0.6 at a height of 150 km and declines down to 0.25 at a height of 250 km.

8.3 Satellites

For a general introduction to space technology several textbooks are available[6].

8.3.1 Solar Panels

A solar panel is a collection of solar cells that convert solar light into electricity (photovoltaics). Lots of small solar cells spread over a large area can work together to provide enough power for satellites or space stations. The more light that hits a cell, the more electricity it produces, so spacecrafts are usually equipped with solar panels that can always be pointed at the Sun even as the rest of the body of the spacecraft moves around.

On Earth, the largest photovoltaic plant is a 10 MW peak power station at Pocking, Germany consisting of 57 912 solar modules delivering 11500 MWh per year. Outside the Earth's atmosphere 1366 W/m^2 are received from the Sun (normal incidence). The atmosphere reflects 6% and absorbs 16% of incoming radiation. The peak power at sea level (1020 W/m^2) may be further reduced by clouds (on the average 20% due to reflection) and absorbtion (16%). Satellite measurements shows that For example, in North America the average power of the solar radiation lies somewhere between 125 and 375 W/m^2(i.e. between 3 and 9 kWh/m^2/day). Currently photovoltaic panels have an efficiency of about 15%, a solar panel delivers 19 to 56 W/m^2 or 0.45-1.35 kWh/m^2/day (annual day and night average).

The most efficient solar panels are the DS1 solar panels which convert about 22 % of the available energy into electrical power. It is also important to note that solar panels lose about 1-2 % of their effectiveness per year. This means after a five year mission, the solar panels will still be making more than 90 % of what they made at the beginning of the mission. Of course this also depends on their distance to the Sun.

There are two major dangers to solar panels in space besides regular wear-and-tear:

- Solar flares that can damage the electronics inside the panels.
- Micrometeorites, which are tiny, gravel-sized bits of rock and other space junk floating in space can scratch or crack solar panels.

Some protection can be made by the use of a thick layer of glass. Of course, if a satellite's mission path takes it away from the Sun (further out into the solar system) solar panels will become less and less efficient.

[6]Gatland K, "Space Technology", Salamander Books (London, 1981)

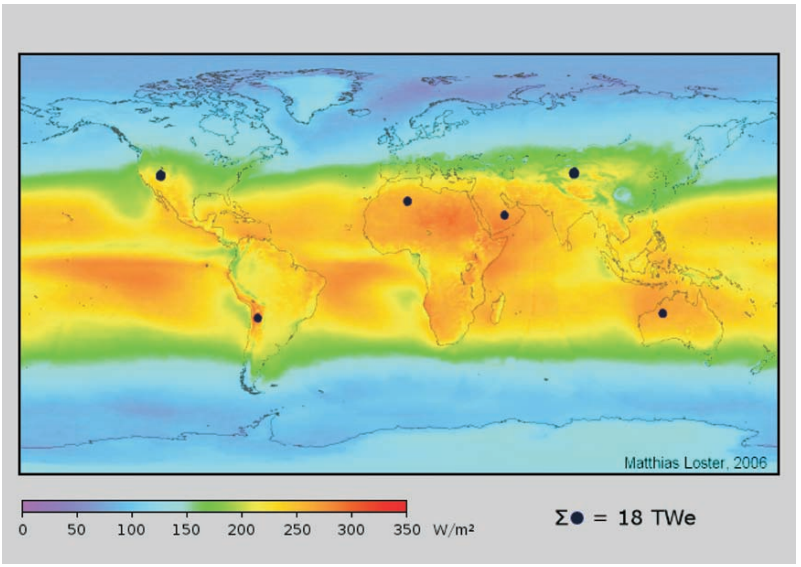


Figure 8.12: Solar power systems installed in the areas defined by the dark disks could provide a little more than the world's current total primary energy demand (assuming a conversion efficiency of 8%). That is, all energy currently consumed, including heat, electricity, fossil fuels, etc., would be produced in the form of electricity by solar cells. The colors in the map show the local solar irradiance averaged over three years from 1991 to 1993 (24 hours a day) taking into account the cloud coverage available from weather satellites. The average electric output would be 18 TW. After: *http* : *//www.ez*2*c.de/ml/solar_land_area/*

Another kind of protection to the above mentioned damaging effects can be made by the use of Fresnel lenses which collect a large area of sunlight and direct it towards a specific spot by bending the rays of light and focussing them- the same principle when people use a magnifying lens to focus the Sunlight on a piece of paper which starts a small fire. Fresnel lenses have been invented in 1822 by Jean Fresnel. Theaters use them for spotlights. They are shaped like a dart board with concentric rings around a lens that is a magnifying glass. Solar concentrators put one of these lenses on top of every solar cell. The solar cells can then be spaced farther apart since the light is focused on each cell. Fewer cells need to be placed and the panels cost less to construct. Thick glass or plastic cover over the solar panel are used to protect them from micrometeorites.

DS1's photovoltaics are made out of gallium arsenide (GaAs). GaAs is made into a cylinder that is then sliced into cells. These solar cells are then connected to the rest of the power network. Solar concentrators, made of clear plastic, are placed above them to focus the Sun's rays.

As a summary we give some literature, further references can be found therein. Markvart *et al.* (1982 [208]) studied the photon and electron degradation of boron-doped FZ silicon solar cells. Radiation-resistant silicon solar cell were investigated

by Markvart *et al.* (1987 [210]). Defect interactions in silicon solar cells were analyzed by Markvart *et al.* (1989 [209]). A study of radiation-induced defects in silicon solar cells showing improved radiation resistance was made by Peters *et al.* (1992 [245]). General information about solar cells can be found in Tada *et al.* (1982 [307]).

A review on radiation damage in solar cells was given by Markvart (1990 [207]).

An analytical study has been carried out on an impact feature within a solar cell from the Hubble Space telescope Solar array. The feature was investigated optically, and the damage was seen as the result of a partially penetrating impact and therefore some impact particles must have been responsible for that. The residue in the impact was found to contain elements such as Fe, Ti, K, Ca, Si, Mg and Na. The elements Mg, Fe and Ti are usually foreign to a solar cell and this suggests that the impact residue may be of natural or man made origin. Subsequent detailed analysis showed Fe and Mg in concentrations of about 10% and Ti in only limited amounts. That implies that the residue is of natural origin. A more detailed description can be found in Graham *et al.* (1997 [120])

8.3.2 Power Sources for Spacecraft

Every power source available for a satellite or other spacecrafts has different strengths and weaknesses. By combining different power sources one can reach an optimum in power generation.

- Batteries: a reliable, well understood technology. However, power demands for satellites tend to be very high and a battery that would be strong enough to power a satellite for the length of a mission would be larger than the satellite itself. Thus, batteries are used as a temporary storage for power from another source.

  A battery can convert chemical energy to electricity by putting certain chemicals in contact with each other in a specific way. Electrons will travel from one kind of chemical to another creating an electric current.

  Batteries come in several styles and NASA spacecraft usually use rechargeable nickel-cadmium or nickel-hydride batteries like those found in laptop computers or cellular phones (DS1 uses nickel-hydrogen batteries).

  Batteries tend to expend their charge fairly quickly. DS1 can last from half an hour to three hours running purely on battery power before the batteries need to be recharged from the solar panels. These batteries are recharged thousands of times during the life of the spacecraft.

- Solar panels: they provide abundant power for nearly all a satellite's needs and are safe and clean to launch. However:

  – solar panels are large and fragile constructions that are vulnerable to damage from external forces or even mechanical failures;

  – they are rather expensive to build and put into space;

Table 8.11: Fuels for RTG's

Element	Half life (years)	Watts/g (thermal)	Watt (thermal)
^{210}Po	0.378	141	570
^{238}Pu	86.8	0.55	3000
^{144}Cs	0.781	25	15
^{190}Sr	28.0	0.93	250
^{242}Cm	0.445	120	495

  - they always need to be pointed at the Sun (not being blocked by planets or other objects);
  - the farther the satellite gets from the Sun, the less effective solar panels work. As a rule of thumb we can state that solar powered missions cannot travel further than the orbit of Mars.

- Radioisotope thermoelectric generators: They are also reliable but tend to be expensive to build and of course there is a risk that radioactive material is set into the environment during a launch failure.

  A radioisotope thermoelectric generator, or RTG, uses the fact that radioactive materials (such as plutonium) generate heat as they decay. The heat is converted into electricity by an array of thermocouples which then power the spacecraft.

  A thermocouple is a device which converts thermal energy directly into electrical energy. Basically, it is made of two kinds of metal that can both conduct electricity. They are connected to each other in a closed loop. If the two metals are at different temperatures, an electric potential will exist between them. When an electric potential occurs, electrons will start to flow, making electric current.

  Another process which belongs to this group of energy generation is nuclear fission where unstable radioactive materials are split into smaller parts. Very large amounts of heat are generated but the whole process is more complex and not as reliable as using the heat produced by radioactive decay. An RTG is steadier.

  Plutonium is a very toxic heavy metal. If it is powdered and inhaled, it is a cancer causing agent. It is sealed inside a hard, radiation proof shell. The shell is designed to survive all conceivable accidents, so even in the unlikely event of a launch failure, none of the radioactive particles should escape.

- Fuel cells: they are similar like batteries but they have a longer lifespan and can be refuelled. They are already in use in the Space Shuttle. However they run hot (400-800^0 C) and the waste heat is often hard to manage.

  When atoms of the two gases oxygen and hydrogen are put next to another, they spontaneously combine to form water. This results in the release of a

lot of energy. In a fuel cell the H and O are separated by a membrane. The refuelling means just to provide more H and O and the waste is pure water. With an external source such as a solar panel, one can split the waste water back into its component parts and use it again as fuel. Fuel cells were first used by the Apollo missions since they last longer than traditional batteries and didn't have expensive radioactive parts.

It is extremely important to control the heat on and around a space ship. The operating temperature is usually given between two numbers like $-10°C$ to $60°C$. The parts of the spacecraft have been tested and will work if the temperature in the spacecraft is between these two numbers. Why does a spacecraft have an operating temperature? For example the rocket thruster can use hydrazine as rocket fuel. Therefore the tanks, plumbing and pumps must be kept at a certain temperature: Hydrazine freezes at $2°C$ and boils at $113°C$. Most electronic components will work only within a narrow range of temperatures, usually $-50°C$ to $+150°C$, components will stop working and make the spacecraft useless if the spacecraft temperatures become too extreme.

Heat tends to expand material parts and the opposite happens when a part is cooled. This problem occurs when one part of the spacecraft is pointed at the Sun and the other one is pointed at empty space. The Sun then heats up only one part and this uneven heating causes the spacecraft to be warped or even break or instruments can be distorted. Another source of heating is caused by electronic components. Heat also makes the electrical system less efficient. Electricity is caused by the flow of electrons and the resistance grows with temperature.

Heat sources can be external (from outside the spacecraft) or internal (from inside the spacecraft). External heat sources include:

- the Sun,
- reflected sunlight from planets and moons,
- heating by friction when travelling through an atmosphere or gas clouds,
- released heat from planets.

Internal heat is generated by the craft's propulsion or electrical system.

8.3.3 Electron Damage to Satellites

Explosive Solar Particle events (SPE) are usually associated with solar flares and coronal mass ejections. Protons and electrons are emitted at high velocities which can cause problems in orbiting satellites. In January 1994 three geostationary satellites suffered failures of their momentum wheel control circuity. One of these satellites never fully recovered. During that period however, no SPE was observed. One explanation for this failure is done by assuming a long duration of high energy electron fluxes that occur during times of high speed solar wind streams. It is important to note that these occur during times of sunspot minimum. Thus not only the electron intensity but the total integrated electron flux is important. The

USAF uses empirically defined values to issue warnings for satellite operators. Damaging conditions are assumed when the daily electron flux (which is given by the number of high energy electrons (> 2MeV) per cm^2 per sterad per day meets either of the following conditions[7]

- greater than 3×10^8 per day for 3 consecutive days; or
- greater than 10^9 for a single day.

Such conditions often occur about 2 days after the onset of a large geomagnetic storm.

How can we determine the probability that surface charging may occur. This can be done by the K-index which, as we have shown in the previous chapter, is a measure for geomagnetic storms. The values of K (3 hourly measure) range from 0-9.

- K=0: quiet
- K$\geq$ 4 surface charges effects could begin,
- K$\geq$ 6 surface charging is probable.

Whereas surface charging usually does not cause big problems, particles with $\geq$1 MeV cause Deep Dielectric Charging. When there occurs a high-speed solar wind stream these particles are concentrated in the Van Allen belts. High energy electrons penetrate the spacecraft's outer surface; they penetrate the dielectric materials such as circuit boards and the insulation in coaxial cables. This gives rise to intense electric fields; as soon as they exceed the breakdown potential of the material they produce sudden discharges (similar to a stroke). This discharge damages the system: components may start to burn, semiconductors may be destroyed. These dielectric charging can be avoided by a special construction of the relevant parts however this leads to additional weight and complexity of the system.

Again, high fluxes of these electrons vary with the 11 year solar cycle and are most prevalent late in the cycle and at solar minimum. The GOES GEO spacecraft measures electron fluxes in the range of 0.6 - 2 MeV (see Fig. 8.13).

8.3.4 Single Event Upsets

Single-event upsets (SEUs) are random errors in semiconductor memory that occur at a much higher rate in space than on the ground. They are non-destructive, but can cause a loss of data if left uncorrected. SEUs are often associated with heavy ions from the galactic cosmic radiation.

What is the cause of SEUs? Energetic charged particles pass through sensitive regions of a chip. Depending on their energy and angle of impact, individual particles can cause a large current impulse sufficient to change the state of a bistable circuit element.

[7] see also *http* : *//www.ips.gov.au/Educational/*1/3/7 :

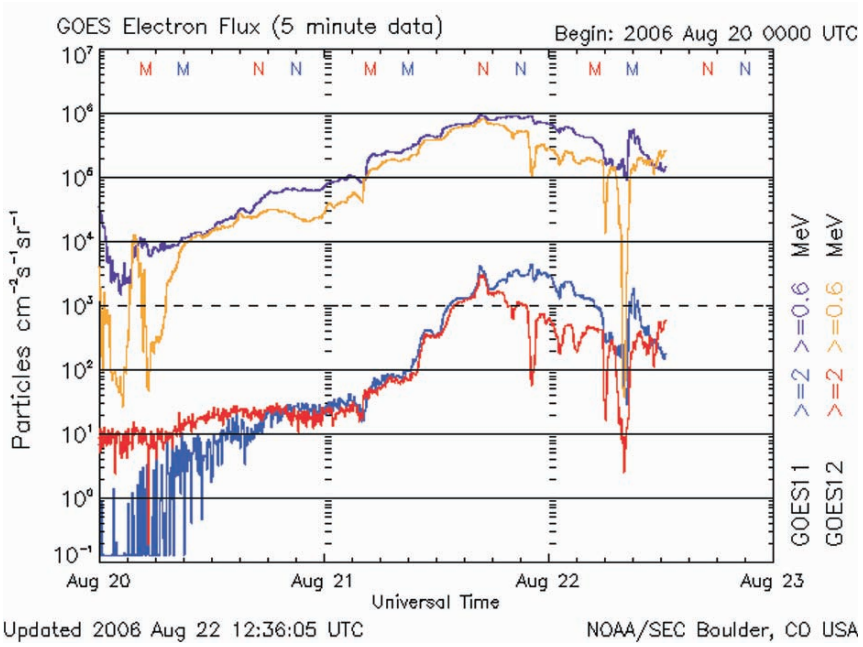


Figure 8.13: Flux measurements by the GOES satellite in different energy channels. *http : //www.sec.noaa.gov/rt_plots/elec_3d.html*

Heavy Ion SEUs occur directly when a heavy ion passes through a semiconductor memory element. The standard models take into account the size, shape, and charge sensitivity of the memory element and the energy, angle, and impact parameter of the incident particle.

For satellites around the Earth, the offset and tilt of the geomagnetic axis with respect to the Earth's rotation axis produces a corresponding miss-alignment of the radiation belts. The result is the South Atlantic Anomaly. The Earth's surface magnetic field is weakest there. Particles drifting around the Earth travel much closer to the Earth than at other latitudes and longitudes.

This higher particle concentration causes a maximum of the distribution of errors in the Atlantic ocean east of the southern part of South America. There occurs also a significant number of errors at high latitudes due to cosmic rays (see Fig. 8.14). These data are from UoSAT-2 which measured from September 1988 to May 1992; UoSAT-2 monitored almost 9000 Single Event Upsets (SEU), and the majority of these (75%) occurred in the South Atlantic Anomaly (SAA) region.

Single event upsets pose also problems to space missions: As a result of volcanic action on Io, the innermost of the large Galilean moons of Jupiter, particles (actually heavy ions) of sulphur and oxygen are present in the space surrounding the planet. These particles form a part of the Jovian magnetosphere. Although

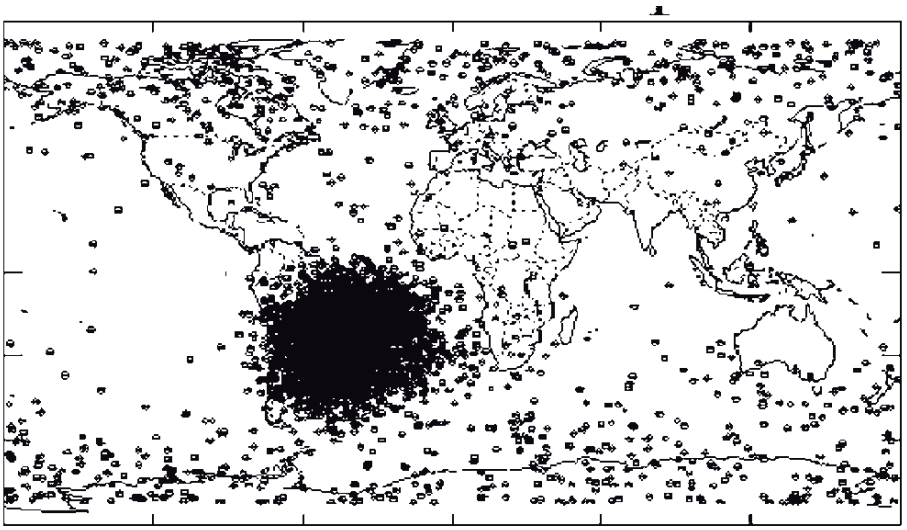


Figure 8.14: Single event upsets; spatial distribution of errors from the UoSAT-3 spacecraft in polar orbit; please note the South Atlantic Anomaly. Adapted from C. Dyer and D. Rodgers, 1998, Space Dep. DERA

the origin of these particles is the moon Io, the volcanoes provide enough velocity for them to escape from the gravitational field of the moon and to become elements of the magnetosphere around Jupiter.

The heavy ions diffuse both inward and outward from the planet. Many of the particles diffuse outward to 20 to 50 times the radius of Jupiter (R_J, measured from the planet's center), where they are accelerated by an interaction with the massive Jovian magnetic field.

The most critical phase of mission operations for to study the Galilean satellites of Jupiter occurs at the time of the spacecraft's closest approach to Jupiter (4 R_J). Heavy ions are capable of penetrating the delicate electronics in the spacecraft and causing a stored computer bit to change its value from a "0" to a "1" or vice-versa, a Single Event Upset results (SEU). A single bit flip in one of Galileo's computer memories could trigger a chain reaction of erroneous commands with disastrous results.

Modern microelectronic devices can suffer from single event effects caused by cosmic radiation neutrons in the atmosphere. The phenomenon has been observed both on ground and at aircraft altitudes. The neutron flux at aircraft altitudes (<15 km) is large enough to make the neutron single event effects a problem to aircraft electronics. The most studied device type is static random access memories (SRAM) since those devices have a very high density of transistors, making them sensitive to particle radiation. The cosmic ray neutrons are produced by the charged primary cosmic radiation in the earth's atmosphere. Thereby the atmospheric neutron flux is certainly influenced by solar activity and space weather

(see e.g. Dyer, 2001 [85])

Normand (1996 [235]) studied the effect of SEU in avionics. Ziegler and Lanford (1979 [346]) studied the effect of cosmic rays on computer memories. SEU in implantable cardioverter defibrillators were studied by Bradley and Normand (1998 [48]). They found some correlation with the expected geographical variation of the secondary cosmic ray flux.

Reedy (1997 [253]) discusses the natural sources of energetic particles in space. The main radiation threats are the galactic cosmic rays (GCRs), solar energetic particles (SEPs) and trapped radiation around planets. Especially outside the Earth's strong magnetosphere, the SEPs are very serious sources of radiation. Over a short period of time (few days) the effects of a huge solar particle event (SPE) ca be greater than any other source of radiation. This causes high doses to humans, microelectronics and solar panels.

Prediction of times with increased risk of internal charging on spacecraft are given by Andersson *et al.* (1999 [9]) and Wu *et al.* (1999 [341]).

8.3.5 Solar Activity and Satellite Lifetimes

Satellites in low Earth orbit, with perigee altitudes below 2000 km, are subject to atmospheric drag. This force very slowly circularizes the orbits and the altitude is reduced too. The rate of decay of these orbits becomes extremely rapid at altitudes less than 200 km. As soon as the satellite is down to 180 km it will only have a few hours to live and after several revolutions around the Earth it will re-entry down to Earth. At that phase the temperature is very high and most of the satellite will vaporize. Only large satellites become not fully vaporized and component pieces of them may reach the ground.

The essential parameter for this deceleration is the air density. This varies along the satellite's orbit and is a function of latitude, longitude, time of day, solar activity etc. At a fixed point in space the density can be expressed in terms of the two space environmental parameters:

- 10 cm solar radio flux (F10),
- geomagnetic index A_p.

It is extremely difficult to predict exactly when a satellite will re-enter the atmosphere[8] . The reason for that is that the space environment is not exactly predictable and there are also unresolved variations in atmospheric density. The accuracy of the prediction is in the order of 10 %. That means that one day before re-entry the uncertainty is at least 2 hours. Within that time however, the satellite will have circled the globe and thus it is difficult to predict the location of re-entry with a reasonable warning.

The decay of a satellite's orbit also depends on the cross section of the object itself. In Fig. 8.15 a rough estimate of the lifetime of a satellite with effective mass

[8]see also: Tobiska W K, R D Culp and C A Barth, "Predicted Solar Cycle Twenty-Two 10.7 cm Flux and Satellite Orbit Decay", Journal of the Astronautical Sciences, 1987, pp419-433, vol 35

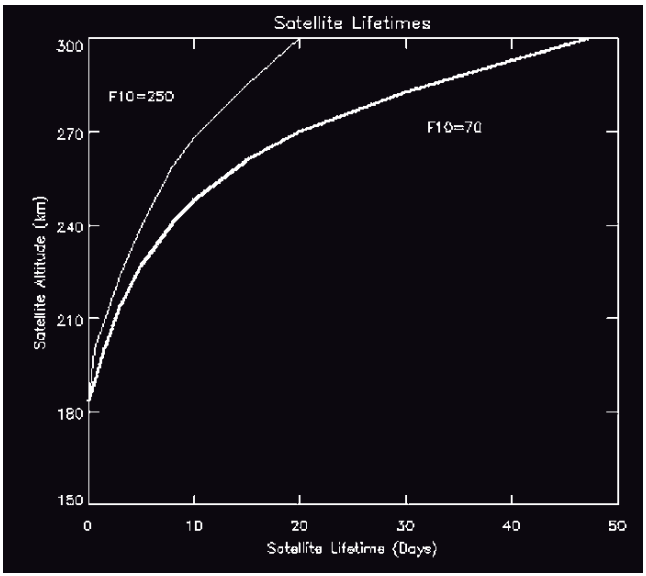


Figure 8.15: Satellite lifetimes

to cross section ratio 100 kg /m^2 in a circular orbit below 300 km is given for two cases: a) for solar minimum conditions, b) for solar maximum conditions. The geomagnetic field is assumed to be quiet during this period. The lifetime values may be varied for satellites of differing mass to area ratios.

The uncertainty in the predictions is shown by a NORAD prediction in April 1979 for the expected re-entry of the SKYLAB space station between 11 June and 1 July of that year. The actual re-entry occurred on July 11, outside the stated interval, a prediction error from mid-interval of around 15%.

8.3.6 Case Study: KOMPSAT1

Here we present a case study: atmospheric drag effects on KOMPSAT1 during geomagnetic superstorms that occurred in Oct 29-30 and Nov. 20, 2003. A satellite travelling through upper atmosphere suffers drag that acts opposite to its orbital motion. This effects low Earth orbiting satellites, LEOs. The atmospheric drag force depends on the satellite's velocity V, the satellite's cross sectional area perpendicular to the direction of motion A, the atmospheric total mass density ρ and the dimensionless drag coefficient C_d:

$$F_d = \frac{1}{2} C_d A V^2 \rho \tag{8.21}$$

The atmospheric density ρ depends on two factors:

- solar EUV radiation, this leads to a heating of the upper atmosphere

- Joule heating: during strong geomagnetic disturbances the contributions to atmospheric heating could reach up to 136% in response to an increase of the K index (Knipp *et al.* 2004 [169]).

The heating of the upper atmosphere can be attributed to different effects. When averaging over longer periods comprising solar maxima and minima the following picture is obtained:

- particles: 36 GW
- Joule heating: 95 GW
- solar: 464 GW (EUV radiation)

These values were obtained for the period 1975 to 2003 (see Knipp *et al*, 2004 [169]). Thus solar wind driven geomagnetic power is about 1/5 of the total budget but becomes dominant when solar activity rises.

The Korean satellite KOMPSAT1 has an altitude of 685 km and was observed for its daily drag acceleration which was found to correlated strongly with geomagnetic disturbances[9]. The results are shown in Fig. 8.14.

8.3.7 The Atmospheric Model

Let us briefly discuss an atmospheric model[10] that is confined to satellites with orbits totally below 500 km altitude. This is an extraction from the decay calculations given at

http : *//www.ips.gov.au/Category/Educational/Space/* .

The reason for that is a simplification: the orbit must then be essentially circular and in place of the orbital radius we can use just the semimajor axis. The atmospheric density ρ is defined by an exponential with variable scale height H. For a fixed exospheric temperature T, H varies with altitude h trough the use of an effective molecular mass m. m includes the actual variation in molecular mass with height and a compensation term for the variation in temperature over the considered range from 180 to 500 km. The variation in density due to the space environment is introduced through T, where $T = T(\mathrm{F10.7\,cm}, \mathrm{A_p})$. Generally, the solar X-ray output incident upon the Earth is absorbed at the base of the thermosphere (120 km) and gives rise to a heating which propagates upward from this level. We use the solar F10.7 cm flux which can vary from 65 to 300 SFU (Solar Flux Units, 1 SFU= $10^{-22}\,\mathrm{W/m^2/Hz}$) as a proxy for X-rays. The other quantity to take into account is the precipitation of particles- most of them coming from the Sun (CMEs). These are well correlated with the large variations in the geomagnetic field measured at the ground level and quantified by the geomagnetic indices (we use the A_p index here). The A_p is computed every 24 hours and during quiet periods just above zero but may rise up to 400.

[9] Kim, K.H., Moon, Y.J., Cho, K.S., Kim, H.D., Park, J.Y., Earth, Planets, Space, 58, e25-e28,2006

[10] see also: King-Hele D, "Satellite Orbits in an Atmosphere - Theory and Applications", Blackie, Glasgow (1987)

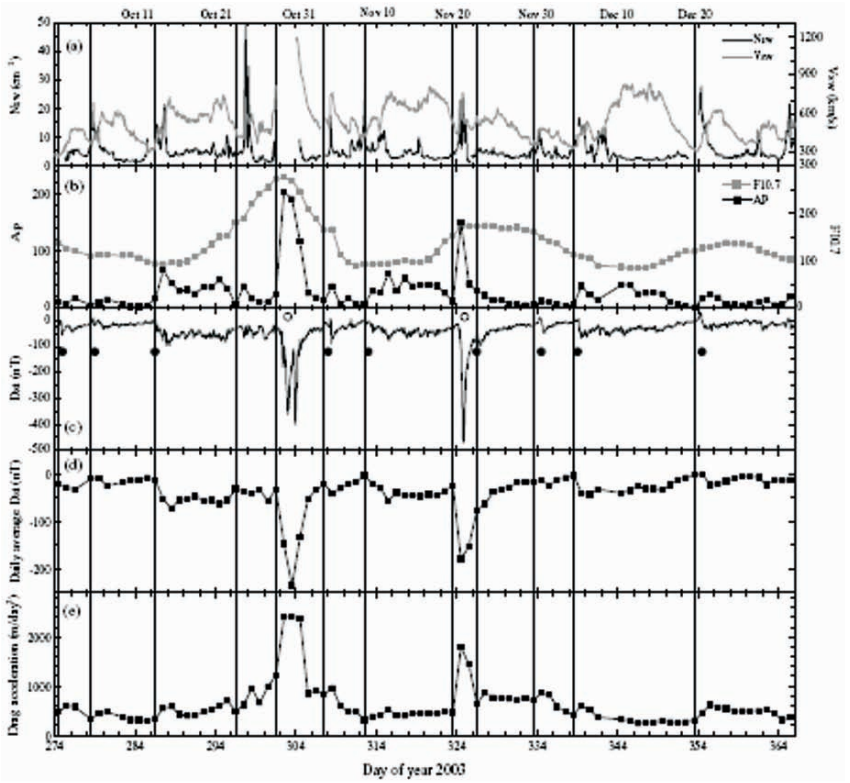


Figure 8.16: Solar wind speed and density from the ACE spacecraft (a), Ap and F10.7 indices (b), provisional Dst index (c), daily averaged Dst (d) and daily KOMPSAT1 acceleration (e). Adapted from http://www.terrapub.co.jp/journals/EPS/pdf/2006e/5806e025.pdf

Then we can write the following set of equations:

$$T = 900 + 2.5(F10.7 - 70) + 1.5A_p \qquad \text{[Kelvin]} \tag{8.22}$$

$$m = 27 - 0.012(h - 200) \qquad 180 < \text{h[km]} < 500 \tag{8.23}$$

$$H = T/m \qquad \text{[km]} \tag{8.24}$$

$$\rho = 6 \times 10^{-10} \exp(-(h - 175)/H) \qquad [\text{kgm}^{-3}] \tag{8.25}$$

The output of this simple model is the density. The intermediate values are only used to derive this density and may not correspond to true atmospheric values at any height within the considered range. The temperature e.g. may be regarded as the mean asymptotic value for the exosphere at large altitudes. The mean molecular weight might be regarded as an integrated mean value from the base of the thermosphere up to the specified height.

The solar 10.7 cm radio flux is used in averaged form (average over the last 90 days). A small correction may be made to weight the current flux more strongly.

Now let us consider the satellite drag. When a spacecraft travels through an atmosphere it experiences a drag force opposite to the direction of its motion. This is given by:

$$D = \frac{1}{2}\rho v^2 A C_d \tag{8.26}$$

D... drag force, ρ atmospheric density, v... speed of the satellite, A... cross sectional area perpendicular to the direction of motion, C_d... drag coefficients. The latter can vary; at altitudes at which satellites orbit $C_d \sim 2$. We introduce the effective cross sectional area $A_e = AC_d$.

For a circular orbit we have the following relation:

$$P^2 G M_e = 4\pi^2 a^3 \tag{8.27}$$

G... gravitational constant, M_e... mass of the Earth. The reduction in the period due to atmospheric drag is given by:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -3\pi a \rho \frac{A_e}{m} \tag{8.28}$$

Re-entry is assumed when the satellite has descended to an altitude of 180 km. The space environmental parameters are given by the solar 10.7 cm radio flux and the geomagnetic activity index. Furthermore one has to provide an estimate for the satellite mass to area ratio. In the absence of any further information this value can be taken as 100 kg /cm^2. This is an average value for many satellites.

If the program underestimates the actual decay of the orbit, you must decrease the mass to area ration, in the case of an overestimation the ratio must be increased.

Also the situation becomes more complicated when considering satellites with very elliptical orbits. Here, a part of the orbit is outside the current atmospheric model. They are also subject to other perturbations (Sun, Moon). If the eccentricity is not too large, one can introduce an effective height in that model:

$$h_{\mathrm{eff}} = q + 900 \times \exp^{0.6} \tag{8.29}$$

q... is the perigee (lowest height) of the orbit and e the eccentricity. For example, the lifetime of a satellite in an elliptical orbit with $e = 0.01$ and $q = 400\,\mathrm{km}$ is the same as the lifetime of a satellite in a circular orbit of height:

$$400 + 900(0.01)^{0.6} \tag{8.30}$$

Since that formula is only a rough approximation it should only be used for orbits with $e < 0.1$. The solar activity should be constant during the orbit decay. Most satellites reaching the end of their lives will have orbits with very low eccentricities (i.e. nearly circular). The reason for this is that atmospheric drag acts to circularize orbits. The apogee height is decreased whilst the perigee height is little affected until the orbit becomes close to circular.

Satellites are perturbed to the first order by the Earth's oblateness and atmospheric drag is important since it acts as energy dissipation. Space weather influences on

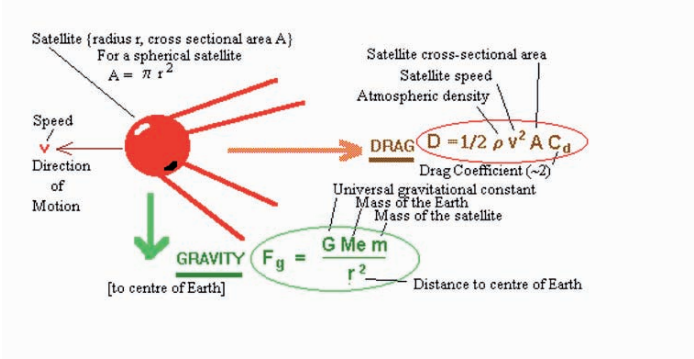


Figure 8.17: Forces acting on a satellite in a low circular orbit. From IPS Radio and Space Services, Satellite Orbital Decay Calculations.

- local air densities
- projected cross section
- local and exospheric temperature
- atmospheric composition

These effects result to varying densities in the atmosphere causing drag effects discussed above. The drivers are the varying solar UV radiation and coronal mass ejections that cause geomagnetic storms. A recent review was given by Doornbos and Klinkrad (2006, [77]).

8.3.8 Special Events

The coupling between geoeffective features in the solar wind and the magnetosphere-ionosphere-thermosphere system is extremely complex. Space weather forecasting requires fast simulations. A complete end-to-end simulation of the October 2003 CMEs leading to one of the most intense solar storms was done by Toth *et al.*, 2005. On October 28, 2003, an X17 flare was observed as well as a full halo coronal mass ejection event. Only 19 hours later, on Oct. 29, 0613 UT, the shock has reached the Earth causing a major geomagnetic storm. The geomagnetic K indices at mid and high latitudes rose up to 9. In Australia[11] HF communication problems were reported, as well as HF communications used by aircraft were disrupted. Aurorae could be seen as far south as Texas. The Japanese satellite Kodama was shut down and damaged. Power grids in the northern US and Canada limited the amount of electricity they produced in order to avoid damage by induced currents. In Malmö, Sweden, a power outage affected 20 000 homes- it is highly possible that this was caused by the event. The crew in the international

[11] see: *http : //www.ips.gov.au/Category/Educational/Space*, vers. june 2006

space station had to move to the aft end of the service module where a higher protection from enhanced radiation was expected.

8.4 Space Weather on Moon and Mars

8.4.1 Spaceweather on Moon

Our Moon is different from the Earth in many ways: it has no atmosphere and no magnetic field. Therefore, the lunar surface is exposed directly to solar radiation and particles. These have eroded the surface layer.

Regolith is a layer of loose, heterogeneous material covering solid rock. On the Moon, regolith has been formed by the action of micro-meteorites breaking down surface rocks into a powder. This powder is more reflective than the basalt that makes up the lunar maria, and therefore looks brighter when viewed from Earth[12].

Meteoroids strike the Moon every day. Lunar escape velocity averages 2.38 km/s. Any rock on the lunar surface that is accelerated by the impact of a meteoroid to lunar escape velocity or greater will leave the Moon's gravitational influence. Some ejected material becomes captured by the Earth's gravitational field and lands on Earth within a few hundred thousands of years or shorter.

The moon does not possess an atmosphere. Meteoroids are not decelerated by such an atmosphere like in the Earth. Lunar regolith is composed in part of rock and mineral fragments that were broken apart from underlying bedrock by the impact of meteorites. A rock composed of bits and pieces of older rocks is called a breccia.

In lunar regolith two types of rocks (called Lithologies) occur:

Agglutinates : Small glassy breccias formed when a micrometeorite strikes the lunar regolith. Micrometeorites have the size of less than 1 mm. When a micrometeorite strikes the lunar surface, some of the impacted regolith melts and some doesn't, so the product is a glass with mineral and rock fragments entrained. The glass often shows flow features. Agglutinates are typically tens of micrometers to a few millimeters in size. All agglutinates contain holes called vesicles. The vesicles all formed from gas bubbles in the glass. Ions from the solar wind (mostly H and He) hit the unprotected surface, not penetrating very deep into the rock (a few 10^{-2} mikrons, thus they are on the surface of the regolith particles). When a micrometeorite strikes fine-grained material at the surface, some of the material gets hot enough to melt and form the glass of an agglutinate. It also gets hot enough to liberate the solar-wind-implanted hydrogen and helium, causing bubbles in the glass.

Glass spherules : These may be formed in two ways. A meteorite impact melts material, the melt is ejected from the crater and small globs of the melt solidify. They are spherical ranging from mm to cm. The other origin is

[12]There were even concerns before the landing of Apollo 11 that the regolith would not be supportive enough to cope with the weight of the lunar module and that the module would begin to sink beneath the surface.

pyroclastic glass, molten rock cools and solidifies above the Moon's surface, leading to glassy spherules.

The regolith contains much information on present and past solar corpuscular radiation. Especially solar wind Ar seems to be well retained in the minerals that were investigated by Wieler *et al.*, 1980 [334] and Wieler *et al.* 1995 [333].

Therefore, the effects of "space weathering" (see also Keller *et al.*, 1999 [158]) on the formation of lunar soils provides the ground-truth requisite to understanding regolith development on all atmosphereless bodies in the solar system, e.g. like asteroids or satellites like Phobos or Deimos. Exposure to the solar wind and solar cosmic rays for long periods of time may cause substantial alterations in the geochemistry of the target material. Experimental results already made in 1967 (see Zeller and Ronca, 1967 [345]) indicate that hydration of oxygen-rich materials can be expected in any surfaces which are exposed to solar protons.

8.4.2 Record of Early Earth Evolution

At this point it is interesting to point out that in lunar even terrestrial components can be found (see also Ozima *et al.*, 2005 [241]. As we have seen in the previous chapters, at present the interaction between the solar wind and the earth's ionosphere is very low because of the shielding of the geomagnetic field. However, if the geomagnetic field is absent, solar wind particles directly can interact with the ionosphere.

Light elements such as N and noble gases in lunar soils are surface correlated and must have been implanted from outside. Most of the N and some of the other volatile elements in lunar soils are actually from the Earth's atmosphere rather than the solar wind (Ozima, 2005 [240]). The existence of terrestrial compounds in some lunar soils suggests that at the time when these components were implanted in Lunar soil, the geomagnetic field was very weak. For example N and ^{26}Ar components can be attributed to this so called earth wind. Moreover one has to take into account that, as was also shown previously, the early Sun was much more active than at present.

8.4.3 Mars

Because Mars will be a target of future manned space missions we briefly discuss space weather influences there.

Mars is a dry like a desert, cold as the Earth's Antarctic and possibly lifeless and future human colonists will be exposed to extreme sets of weather conditions. The Earth is protected by the magnetosphere. Mars does not possess a global magnetic field to shield the surface from SEPs and cosmic rays. The solar wind gradually eroded the martian atmosphere and at present the surface pressure is 1-9 millibars, depending on altitude; the average is 7 mb.

In astrobiology it is believed that protection by an atmosphere and magnetic field are essential factors for life on a planet.Does that mean that because Mars has no global magnetic field and a very thin atmosphere the planet is lifeless?

It can be assumed that certain life forms could be radiation resistant like the terrestrial microbe Deinococcus radiodurans. D. radiodurans has a feature that is considered all-important in aerospace: redundancy. Its genetic code repeats itself many times so that damage in one area can be recognized and repaired quickly. It withstands attacks from acid baths, high and low temperatures, and even radiation doses, e.g. the microbe can withstand without loss of viability a dosage that is 3 000 times greater than what would kill a human.

Tiny Martians might also live in rocks or soil, substances that provide natural protection against radiation. In March 2004, small quantities of methane (about 10 parts in a thousand million) were detected in the Martian atmosphere by researchers operating the Planetary Fourier Spectrometer (FPS) experiment on the orbiting Mars Express spacecraft and also by astronomers using the Keck Telescopes and the Canada-France-Hawaii Telescope on Earth (see e.g. Krasnopolsky, 2006 [172]). Normally, under conditions in the Martian atmosphere Methane should be destroyed by UV solar radiation within a few hundred years. Its presence can be explained either by volcanic outgassing or by the action of methanogens (bacteria that produce methane).

As we have seen in the previous chapter, particulate radiation poses the greater threat to humans. Particles from solar flares (protons) are of greater concern - here particles with relatively low energies (around 70 MeV) are produced. Such protons lose energy in tissue. Cosmic ray nuclei have energies between 300 and 500 MeV and they penetrate the human body in a short time so that there is not enough time to transfer their energy to the surrounding tissue. Solar protons when passing through humans ionize molecules along their track- the ionization creates free radicals causing modification or even break of the DNA strands and if the cell survives it can become cancerous.

Therefore, human settlers on Mars must be protected from these energetic protons. The air density at the surface of Mars is equivalent to that of the Earth's atmosphere at 20 km altitude. Astronauts must be protected by shelter walls.

Another problem on Mars which is not directly related to space weather is martian dust. Mars is red because its surface is largely composed of iron oxide (rust) and oxides of other minerals. Some scientists suspect that the dusty soil on Mars may be such a strong oxidizer that it burns any organic compound such as plastics, rubber or human skin as viciously as undiluted lye or laundry bleach. Data from the Pathfinder mission showed that Martian dust may also contain trace amounts of toxic metals, including arsenic and hexavalent chromium.

Since Mars has no global magnetic field the surface is eroded by the solar wind as well as by the planet's atmosphere. Currently Mars looses approximately 2 kg/s of its atmosphere. In 1998 magnetometers discovered a network of magnetic loops arrayed across Mars's southern hemisphere (see Fig 8.18). Locally, the magnetic fields arch over the surface like umbrellas, hundreds of km high. In such an area you would measure a field about as strong as the Earth's (a few tenths of a gauss). Elsewhere the field is extremely weak. The martian ionosphere traces the distribution of the surface magnetic field, and there seems to be a 1-to-1 correspondence: places where magnetic umbrellas deflect the solar wind are also spots where the ionosphere is retained high above the surface of the planet.

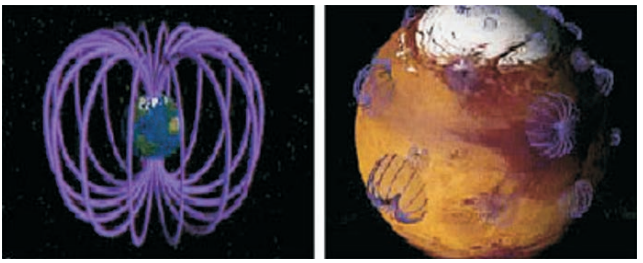


Figure 8.18: Artistic illustration of Earth magnetic field and Mars magnetic field (not influenced by the solar wind). Credit: NASA

The Earth's global magnetic field is caused and maintained from an active dynamo – that is, circulating currents at the planet's liquid metallic core. A similar dynamo once churned inside Mars, but for reasons unknown it stopped working four billion years ago. The patchwork fields mentioned above, we see now are remnants of that original magnetic field. Hellas and Argyre, two large impact basins on Mars are about four billion years old and are demagnetized. If the dynamo was still operating when those impact features formed, the crust would have re-magnetized as they cooled. Hence, the dynamo must have stopped before then. The strongest magnetic anomalies were found in Terra Sirenum which might therefore be the most suitable landing site for manned missions (Alves and Baptista, 2004[7]).

The Mars Odysssey spacecraft (2001) has onboard the Martian radiation environment experiment (MARIE) which measures the background radiation due to galactic cosmic rays and solar protons. Good agreement between models and measurements were found (Atwell *et al.* 2004,[16]).

Space Radiation Hazards on Human Missions to the Moon and Mars are described in Townsend, 2004 [314].

Crewmembers of the ISS will be exposed to ionizing radiation (the inclination of ISS to the Earth orbit is 51 degrees, it is in low Earth orbit, LEO). The concept is as low as reasonably achievable (ALARA) radiation exposure and cheap Poylethylene, C_nH_n with low atomic number can be used for shielding (Shavers *et al.*, 2004 [280]). Models like HZETRN allow computer simulations of radiation hazards (Wilson *et al.*, 2004 [339]).

The complex reactions that are caused by solar radiation and particles in different planetary atmospheres are discussed (mainly for the UV radiation) in the book of Vazquez and Hanslmeier, 2006 [325].

Chapter 9

Real-Time Space Weather and Forecasts

In this chapter we will give an overview over existing real-time space weather centers and the data available there. First we will discuss a very useful classification scheme.

9.1 NOAA Space Weather Scales

As we have seen fast and efficient communication of space weather effects to the public is very important. For that reason the US NOAA (National Oceanic and Atmospheric Administration) has introduced the space weather scales. A summary of the different influences triggered by the Sun is shown in Fig. 9.1.

The NOAA space weather scales can be grouped into three different parts:

- Geomagnetic storms
- Solar radiation storms
- Radio blackouts

In the following we will briefly review these scales.

9.1.1 Geomagnetic Storms

The geomagnetic storms are divided into 5 categories, G1...G5 where the last have the most severe effects.

G1

classified as minor; the influence on power systems is weak, some grid fluctuations can occur. Also the influence on spacecraft is negligible. It seems however that

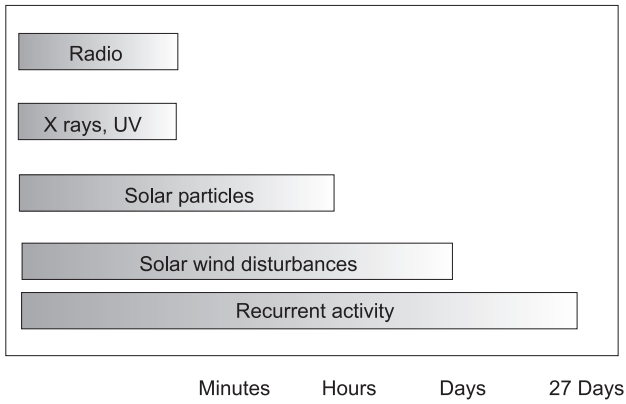


Figure 9.1: After a flare or coronal mass ejection erupts from the Sun's surface, major disturbances arrive with a range of time delays and a storm begins to build in the space surrounding the Earth.

migratory animals are affected even at this low level; the aurora is commonly visible at high latitudes.

As an average about 1700 events per cycle (corresponding to about 900 days per cycle) are to be expected. The K_p value is about 5.

G2

moderate; at this activity some damage may occur in power systems: high-latitude power systems may experience voltage alarms, long-duration storms may cause transformer damage.

Concerning spacecraft operations corrective actions to orientation may be required by ground control; possible changes in drag affect orbit predictions. This imposes problems to fully automated satellites.

Concerning terrestrial telecommunication the HF radio propagation can fade at higher latitudes, and aurora has been seen as low as New York and Idaho (down to 55 geomagnetic latitude). The K_p value is about 6 and on the average one can expect 600 events per cycle (corresponding to about 360 days per cycle).

G3

strong; on power systems voltage corrections may be required; furthermore false alarms can be triggered on some protection devices.

On satellite components surface charging may occur. Due to the extension of the terrestrial atmosphere during these events drag may increase on low-Earth-orbit satellites, and corrections may be needed for orientation problems.

It is also very important to note that intermittent satellite navigation and low-frequency radio navigation problems may occur, HF radio may be intermittent, and aurora has been seen as low as down to 50° geomagnetic lat. The K_p value

is about 7 and on the average one can expect 200 events per cycle (corresponding to 130 days per cycle).

G4

severe; widespread voltage control problems may occur in power systems and some protective systems will mistakenly trip out key assets from the grid.

The problems of surface charging and tracking of satellites increase considerably.

On surface pipelines, induced currents affect preventive measures; the satellite navigation can degrade for hours and low frequency navigation can be disrupted. Aurora has been seen down to 45^0 geomagnetic latitude. The K_p index is at 8 and on the average one has to count with 100 events per cycle (corresponding to 60 days per cycle).

G5

extreme; widespread voltage control problems and protective system problems can occur; transformers may experience damages and some grid systems may experience complete collapse or blackouts.

The spacecraft operations are affected by extensive surface charging, problems with orientation, uplink/downlink and tracking satellites.

At this activity pipeline currents can reach hundreds of amperes, HF radio propagation may be impossible in many areas for one to two days, satellite navigation may be degraded for days, low-frequency radio navigation can be blocked for hours. The aurora has been seen down to 40^0 geomagnetic latitude (Italy, southern Texas). At this level $K_p = 9$ and on the average one has to expect 4 events per cycle (corresponding to about 4 days per cycle).

9.1.2 Solar Radiation Storms

Again this is a classification from 1 to 5 (S1...S5). This activity can be quantitatively measured by the flux of ions $\geq 10\,\mathrm{MeV}$ as five minute averages in the units $\mathrm{s}^{-1}\mathrm{ster}^{-1}\mathrm{cm}^{-2}$.

S1

minor; there are no effects on biological systems and satellite operations; no danger for astronauts (especially for EVAs (extravehicular activities)). There may be some minor impacts on HF radio in the polar regions. The ion flux is about 10 (see above units). There are about 50 events per solar cycle.

S2

moderate; there are no biological influences; for satellite operations it is important to know that infrequent single-event upsets are possible.

Small effects occur on HF propagation through the polar regions and navigation at polar cap locations is possibly affected.

The ion flux is about 100 and can expect about 25 events per cycle.

S3

strong; at this level radiation hazard avoidance is recommended for astronauts on EVA; passengers and crew in commercial jets at high latitudes may receive low-level radiation exposure (equivalent to approximately 1 chest x-ray).

The effects on satellite operations become important: lots of single-event upsets, noise in imaging systems, and slight reduction of efficiency in solar panels are likely.

On Earth, degraded HF radio propagation through the polar regions and navigation position errors are likely. The ion flux is about 10^3 and about 10 such events per cycle may occur.

S4

severe; unavoidable radiation hazard to astronauts on EVA thus it is necessary to alarm astronauts; moreover, elevated radiation exposure to passengers and crew in commercial jets at high latitudes (equivalent to approximately 10 chest x-rays) is possible.

Satellites may experience memory device problems and noise on imaging systems; star-tracker problems may cause orientation problems, and solar panel efficiency can be degraded.

On the surface blackout of HF radio communications through the polar regions and increased navigation errors over several days are likely.

The ion flux is about 10^4. There are about 3 such events per cycle.

S5

extreme; unavoidable high radiation hazard to astronauts on EVA; high radiation exposure to passengers and crew in commercial jets at high latitudes (equivalent to approximately 100 chest x-rays) is possible.

Satellites may be put out of operation, memory impacts can cause loss of control, may cause serious noise in image data, star-trackers may be unable to locate sources; permanent damage to solar panels is possible.

At the surface complete blackout of HF communications is possible through the polar regions, and position errors make navigation operations extremely difficult.

The ion flux is at 10^5; fortunately, these events occur on a rate fewer than 1 per cycle.

9.1.3 Scale for Radio Blackouts

Classified as R1...R5; measured as GOES X-ray peak brightness by class, measured in the 0.1-0.8 nm range, in Wm^{-2}.

R1

minor; we have to take into account a weak or minor degradation of HF radio communication on the sunlit side, as well as occasional loss of radio contact.

Concerning navigation we have to consider that low-frequency navigation signals may be degraded for brief intervals. The flare classification is M1 and (10^{-5}). On the average 2000 such perturbances per cycle occur (on 950 days per cycle).

R2

moderate; limited blackout of HF radio communication on sunlit side occur, loss of radio contact for tens of minutes.

Navigation: a degradation of low-frequency navigation signals for tens of minutes is likely. The classification of the relevant solar event goes M5 and the flux to 5×10^{-5}.

On the average one has 350 events per cycle (300 days per cycle).

R3

strong; a wide area blackout of HF radio communication, as well as a loss of radio contact for about an hour on the sunlit side of Earth is likely.

Since low-frequency navigation signals are being degraded for about an hour this also has serious consequences for navigation.

The classification is X1, the flux 10^{-4} and 175 events per cycle (140 days per cycle) are probable.

R4

severe; HF radio communication blackout occurs mostly on the sunlit side of Earth for one to two hours and a HF radio contact loss during this time has to be expected.

Outages of low-frequency navigation signals cause increased error in positioning of navigational systems for one to two hours. Minor disruptions of satellite navigation are likely on the sunlit side of Earth.

The classification is X10, the flux 10^{-3} and 8 events per cycle (8 days per cycle) can be expected.

R5

extreme; a complete HF (high frequency) radio blackout on the entire sunlit side of the Earth lasting for a number of hours may occur. This results in no HF radio contact with mariners and en route aviators in this sector.

Navigation: Low-frequency navigation signals used by maritime and general aviation systems experience outages on the sunlit side of the Earth for many hours, causing loss in positioning. Satellite navigation errors in positioning increase for several hours on the sunlit side of Earth, which may spread into the night side.

The classification is X20, the flux 2×10^{-3} and there are less than 1 events per cycle.

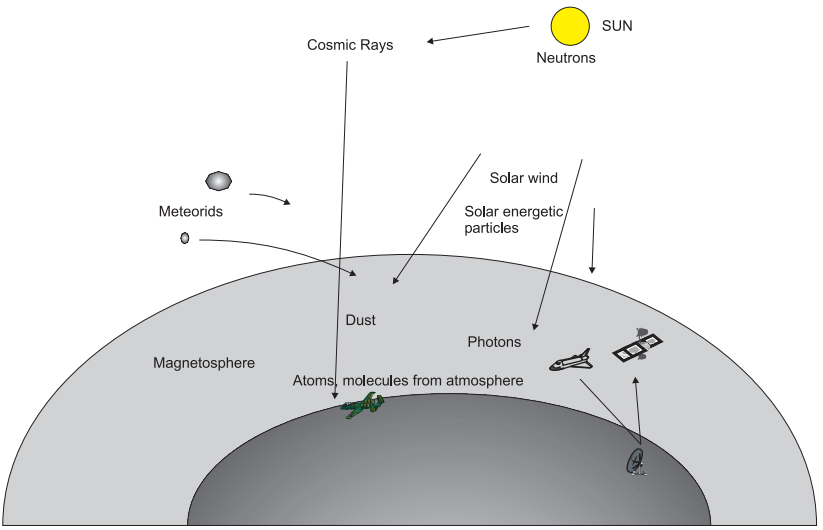


Figure 9.2: Summary of space weather effect in the Earth's environment

9.1.4 Summary

The classification scheme given above enables very easily to estimate the effect of geomagnetic storms and solar radiation storms on satellites and telecommunication systems. This is also extremely important for manned space mission (ISS, international space station). On the other hand, the solar activity has declined after heaving reached its maximum in 2000, the new cycle will start in 2006. One can estimate that there will be about 25 EVA/year necessary for the construction of the space station. For that reason, it is extremely important to alert astronauts for S4 and S5 storms. The predicted sales figures for GPS systems rise from 5000 Million USD for 1998 and more than 9000 Million USD for 2000. This means that more and more systems are equipped with these navigation systems but on the other hand we must take into account that small degradations may even occur at R1 levels. The frequency of such events is however more than 2000 per cycle.

Also the number of satellites will increase.

We have seen before that some military detection or early-warning systems are also affected by solar activity. The Over-the-Horizon Radar bounces signals off the ionosphere in order to monitor the launch of aircraft and missiles from long distances. During geomagnetic storms, this system can be severely hampered by radio clutter. That can occur at even low activity (R1 perturbances).

9.2 The Main Space Weather Sources

Currently the three major sources are:

- NOAA Space Environment Center (Boulder, CO, USA)

- Solar-Terrestrial Dispatch (University of Lethbridge, Stirling, Alberta, Canada)
- Australian Space Forecast Centre (Haymarket, NSW, Australia).

In addition to these sources many other observatories contribute to space weather related data. A list of these institutions and observatories can be found on the homepage of the Swedish Lund Space Weather center
http : *//www.lund.irf.se/*
and at the website
http : *//www.lund.irf.se/HeliosHome/spwfo.html*

9.2.1 NOAA Environment Center

The SEC provides regular bulletins and forecasts as well as solar images and scientific data. Via WWW you can get the latest Geophysical alert message text like:

:Product: Geophysical Alert Message wwv.txt :Issued: 2006 Aug 23 0911 UTC
Prepared by the US Dept. of Commerce, NOAA, Space Environment Center
Geophysical Alert Message
Solar-terrestrial indices for 22 August follow. Solar flux 81 and mid-latitude A-index 21. The mid-latitude K-index at 0900 UTC on 23 August was 3 (31 nT).
No space weather storms were observed for the past 24 hours.
No space weather storms are expected for the next 24 hours.

Furthermore, you can get a Report of Solar and Geomagnetic Activity –issued daily at 2200Z. This summarizes solar and geomagnetic conditions for the past 24 hours and provides forecasts for the next three days and moreover contains that day's Solar Flux value and A-index, plus forecasts for the next three.

Another valuable information is the D-region Absorption Prediction –updated once a minute; this color-coded world map graphic depicts the current position of the sun and how a solar flare might have caused an HF fade.

In *Today's Space Weather* more information including full-disk H-Alpha solar images etc can be found.

9.2.2 Solar-Terrestrial Dispatch

STD focuses great emphasis upon HF communications. Current solar flux value (updated every 30 minutes), Boulder and planetary K-indices (computed each 3 hours) for the current UTC day, running estimated 24-hour planetary A-index, plus K- and A-indices for the previous UTC day can be found there as well as the 10.7 cm solar radio flux.

Another feature are MUF maps. The map shows the radio auroral zones as green bands near the northern and southern poles. The area within the green bands is known as the auroral zone. Radio signals passing through these auroral zones will experience increased signal degradation in the form of fading, multipathing and absorption.

The radio auroral zones are typically displaced equatorward from the optical auroral zones (or the regions where visible auroral activity can be seen with the eye).

A world map showing the highest vertically directed frequency being returned from the ionosphere are the critical F2-layer maps.

9.2.3 Australian Space Forecast Centre

Besides current solar, HF, ionospheric, and geomagnetic conditions e.g. a recent X-ray flare list is available where you get almost instantaneous data on flare activity in the recent past. The data are given as:

Approximate Flare Start : 06-07-2006 0823 UT
Approximate Flare Maximum: 06-07-2006 0837 UT at Flux M 2.5
Approximate Flare End : 06-07-2006 0859 UT
LOCATION OF HF FADEOUT: Middle East

There are also online prediction tools to predict HF propagation paths. Also frequencies and times for HF communication paths between two user-entered points anywhere in the world is provided by GRAFEX. On a world map, you click your location, enter up to ten operating frequencies, the desired UTC hour, and then use the mouse to drag out a target region. Now you choose either real-time or forecast ionospheric conditions and receive a color-coded coverage map for all frequencies plus individual maps for each chosen frequency.

9.3 Space Weather Forecasts

As we have seen in the previous sections, phenomena on the Sun as flares, coronal mass ejections, solar wind, are coupled to the Earth's magnetosphere/atmosphere system. High energetic events on the Sun can lead to a destruction of satellites and therefore it becomes more and more desirable to predict solar activity. We summarize some recent attempts.

In the section about helioseismology we described, how the interior of the Sun can be investigated by analyzing solar eigenmodes. This enabled us to study the structure around and below visible active regions such as flows and temperatures. For local area helioseismology data from SOHO and GONG are extremely important and will hopefully help to improve predictions about solar active phenomena. Physics-based neural network to investigate how the flow fields can be used to predict flares and space weather and to incorporate these data into the space weather forecast models was done by Jensen *et al.* 2004 [148].

A software tool that automatically detects Coronal Mass Ejections as observed by the LASCO C2 instrument on board the SOHO spacecraft was developed by Olmedo *et al.* 2006 [237]. The automatic method gives a true positive rate of approximately 75% using a manual catalog as a benchmark and is part of software tools in the Solar Eruptive Event Detection System (SEEDS).

During the declining phase of solar cycle 23 four epochs of extremely high flare activity occurred. This was not observed in the previous cycles 21 and 22 (only 1 event) but similar to cycle 20 (Bai, 2006 [24]).

The Solar Mass Ejection Imager (SMEI) has been tracking coronal mass ejections (CMEs) from the Sun to the Earth and beyond since February 2003. A prediction of the arrival time at ACE within 2 hours of its actual arrival for three events, and within 10 hours for eight events was possible. Of these eight events, seven were detected by SMEI more than 1 day before the transient's arrival at the Earth (Howard *et al.* 2006 [139]).

The solar cycle 24, that started in 2006 could have a 30-50% higher peak than cycle 23. This prediction is based on a flux transport dynamo model that gave good predictions for the peaks of cycles 16-23 (Dikpati *et al.*, 2006 [76]). This is in contradiction with the prediction given by Schatten, 2005 [272] using the polar field precursor method of solar activity forecasting. In that paper, the peak amplitude of the next solar cycle 24 is estimated at 124 ± 30 in terms of smoothed F10.7 Radio Flux and 80 ± 30 in terms of smoothed international or Zurich Sunspot number. That would mean 'fair space weather'.

Burov, 2006 [51] wrote a paper on the possibility of getting economically sound forecasts of rare space weather events where he mentioned the need for precise forecasting of major events.

Space weather prediction by cosmic rays was investigated by Mavromichalaki *et al.*, 2006 [213] using real-time data from a neutron monitor network. The system collects data in real-time mode from about 15 real-time cosmic ray stations by using the internet. The main server in Athens station collects 5-min and hourly cosmic ray data and produces forecasting.

Medium range (1– 3 days) geomagnetic forecasting using data of slow/high speed solar wind stream interfaces were described by Gleisner and Watermann, 2006 [114].

An approach to space weather via analysis, prediction, modeling, and classification supported by Genetic Programming was made by Jorgensen and Karimabadi (2005 [150]). Genetic programming can search spaces of algorithms or mathematical functions. A population of candidate solutions is allowed to evolve. Then a natural selection and survival of the fittest starts.

Chapter 10

Asteroids, Comets, Meteroites

10.1 Asteroids

10.1.1 General Properties

On the first day of January 1801, Giuseppe Piazzi discovered an object which he first thought was a new comet. But after its orbit was better determined it was clear that it was not a comet but more like a small planet and it was therefore named asteroid. The proper name of the first asteroid detected is Ceres[1]. Three other small bodies were discovered in the next few years (Pallas, Vesta, and Juno). By the end of the 19th century several hundred asteroids were known.

Several thousand asteroids have been discovered and given provisional designations so far. Thousands more are discovered each year. There are 26 known asteroids larger than 200 km in diameter. About 99 % of all objects > 100 km are known however of the total number of asteroids with diameters between 10 and 100 km we know only 50%. It is difficult to estimate the total number of asteroids, perhaps as many as a million 1 km sized asteroids may exist most of them being too small to be seen from the Earth.

Since most of the asteroids have orbits between Jupiter and Mars, it was first assumed that they are remnants of a larger planet that broke up. However, the total mass of all the asteroids is less than that of the Moon[2].

Ceres has a diameter of 933 km, the next largest are Pallas, Vesta and Hygiea which are between 400 and 525 km in diameter. All other known asteroids are less than 340 km.

[1] According to the new definition of the international Astr. Union adopted in 2006, Ceres and Pluto are classified as dwarf planet

[2] The mass of the moon is only 1/81 Earth masses

10.1.2 Classification of Asteroids

Asteroids are classified into:

- C-type: extremely dark (albedo 0.03), similar to carbonaceous chondrite meteorites; 75% of known asteroids belong to this class.
- S-type: 17% of asteroids; bright (albedo 0.1-0.2); metallic Ni, Fe and Mg silicates.
- M-type: bright (albedo 0.1-0.2), pure NiFe.
- rare types

One should however take into account biases in the observations- e.g. dark C-types are more difficult to detect.

According to their position in the solar system, asteroids can also be categorized into:

- Main belt: located between Mars and Jupiter, 2-4 AU from the Sun.
- Near Earth Asteroids (NEAs): they closely approach the Earth and will be treated separately.
- Trojans: located near Jupiter's Lagrange points (60 degrees ahead and behind Jupiter in its orbit); several 100 are known.
- Between the main concentration in the Main Belt are relatively empty regions known as Kirkwood gaps. These are regions were an object's orbital period would be a simple fraction of that of Jupiter (resonance).
- Centaurs: asteroids in the outer solar system; e.g. Chiron (his orbit lies between Saturn and Uranus).

10.2 Impacts by Asteroids

10.2.1 Potentially Hazardous Asteroids

Potentially Hazardous Asteroids (PHAs) are currently defined based on parameters that measure the asteroid's potential to make threatening close approaches to the Earth.

To be classified as PHA, the following parameters must be fulfilled:

- an Earth Minimum Orbit Intersection Distance (MOID) of 0.05 AU or less,
- absolute magnitudes (H) of 22.0 or less are considered.

An asteroid's absolute magnitude H is the visual magnitude an observer would record if the asteroid were placed 1 Astronomical Unit (AU) away, and 1 AU from the Sun and at a zero phase angle. The diameter of an asteroid can estimated from its absolute magnitude (H). The lower the H value, the larger the size of the

object. However, this also requires that the asteroid's albedo be known as well. Since the albedo for most asteroids is not known, an albedo range between 0.25 to 0.05 is usually assumed. This results in a range for the diameter of the asteroid. The table 10.2 shows the diameter ranges for an asteroid based on its absolute magnitude, assuming an albedo ranging from 0.25 to 0.05.

In other words, asteroids that can't get any closer to the Earth (i.e. MOID) than 0.05 AU (roughly 7,480,000 km) or are smaller than about 150 m in diameter (i.e. H = 22.0 with assumed albedo of 13%) are not considered PHAs. The current list of PHAs is obtained from the Minor Planet Center on a daily basis. Asteroids with a small MOID to Earth should be carefully followed because they can become Earth colliders[3].

Because of long-range planetary gravitational perturbations and, particularly, close planetary approaches, asteroid orbits change with time. Consequently, MOID also changes. As a rule of thumb, MOID can change by up to 0.02 AU per century, except for approaches within 1 AU of massive Jupiter, where the change can be larger. Thus, an asteroid that has a small MOID with any planet should be monitored. Currently there are about 350 known PHA's.

10.2.2 Torino Impact Scale

This was established (analogous to the space weather scale) to characterize different objects.

Events Having No Likely Consequences (White Zone)

0 The likelihood of a collision is zero, or well below the chance that a random object of the same size will strike the Earth within the next few decades. This designation also applies to any small object that, in the event of a collision, is unlikely to reach the Earth's surface intact.

Events Meriting Careful Monitoring (Green Zone)

1 The chance of collision is extremely unlikely, about the same as a random object of the same size striking the Earth within the next few decades.

Events Meriting Concern (Yellow Zone)

2 A somewhat close, but not unusual encounter. Collision is very unlikely.
3 A close encounter, with 1% or greater chance of a collision capable of causing localized destruction.
4 A close encounter, with 1% or greater chance of a collision capable of causing regional devastation.

[3]see also: Impacts on Earth, by D. Benest, Springer, 1998

Threatening Events (Orange Zone)

5 A close encounter, with a significant threat of a collision capable of causing regional devastation.
6 A close encounter, with a significant threat of a collision capable of causing a global catastrophe. **7** A close encounter, with an extremely significant threat of a collision capable of causing a global catastrophe.

Certain Collisions (Red Zone)

8 A collision capable of causing localized destruction. Such events occur somewhere on Earth between once per 50 years and once per 1 000 years.
9 A collision capable of causing regional devastation. Such events occur between once per 1 000 years and once per 100 000 years.
10 A collision capable of causing a global climatic catastrophe. Such events occur once per 100 000 years, or less often.

10.2.3 NEOs

Near-Earth Objects (NEOs) are comets and asteroids that have been nudged by the gravitational attraction of nearby planets into orbits that allow them to enter the Earth's neighborhood. Composed mostly of water ice with embedded dust particles, comets originally formed in the cold outer planetary system while most of the rocky asteroids formed in the warmer inner solar system between the orbits of Mars and Jupiter. The scientific interest in comets and asteroids is due largely to their status as the relatively unchanged remnant debris from the solar system formation process some 4.6 billion years ago. The giant outer planets (Jupiter, Saturn, Uranus, and Neptune) formed from an agglomeration of billions of comets and the left over bits and pieces from this formation process are the comets we see today. Likewise, today's asteroids are the bits and pieces left over from the initial agglomeration of the inner planets that include Mercury, Venus, Earth, and Mars.

As the primitive, leftover building blocks of the solar system formation process, comets and asteroids offer clues to the chemical mixture from which the planets formed some 4.6 billion years ago. If we wish to know the composition of the primordial mixture from which the planets formed, then we must determine the chemical constituents of the leftover debris from this formation process - the comets and asteroids.

In terms of orbital elements, NEOs are asteroids and comets with perihelion distance q less than 1.3 AU. Near-Earth Comets (NECs) are further restricted to include only short-period comets (i.e orbital period P less than 200 years). The vast majority of NEOs are asteroids, referred to as Near-Earth Asteroids (NEAs). NEAs are divided into groups (Aten, Apollo, Amor) according to their perihelion distance (q), aphelion distance (Q) and their semi-major axes (a).

Possible NEO missions that require spacecraft with the capability to rendezvous at great distances (1 AU) from the Earth within a releatively short amount of time (on the order of a year) are discussed by Sforza and Remo (1997 [279]) and

Table 10.1: Groups of Asteroids near Earth orbit

Group	Description	Definition
NECs	Near-Earth Comets	q<1.3 AU, P<200 years
NEAs	Near-Earth Asteroids	q< 1.3 AU
Atens	Earth-crossing NEAs	a<1.0 AU,Q>0.983 AU
Apollos	Earth-crossing NEAs	a>1.0 AU, q<1.017 AU
Amors	Earth-appr. NEAs with orbits betw. Earth and Mars	a>1.0 AU, 1.017<q<1.3 AU
PHAs	NEAs	MOID≤0.05 AU, H≤22.0

Table 10.2: Absolute magnitude and diameter of asteroids

absolute Magn.	Diameter	absolute Magn.	Diameter
3.0	670 km - 1490 km	3.5	530 km - 1190 km
4.0	420 km - 940 km	4.5	330 km - 750 km
5.0	270 km - 590 km	5.5	210 km - 470 km
6.0	170 km - 380 km	6.5	130 km - 300 km
7.0	110 km - 240 km	7.5	85 km - 190 km
8.0	65 km - 150 km	8.5	50 km - 120 km
9.0	40 km - 90 km	9.5	35 km - 75 km
10.0	25 km - 60 km	10.5	20 km - 50 km
11.0	15 km - 40 km	11.5	13 km - 30 km
12.0	11 km - 24 km	12.5	8 km - 19 km
13.0	7 km - 15 km	13.5	5 km - 12 km
14.0	4 km - 9 km	14.5	3 km - 7 km
15.0	3 km - 6 km	15.5	2 km - 5 km
16.0	2 km - 4 km	16.5	1 km - 3 km
17.0	1 km - 2 km	17.5	1 km - 2 km
18.0	670 m - 1500 m	18.5	530 m - 1200 m
19.0	420 m - 940 m	19.5	n330 m - 750 m
20.0	270 m - 590 m	20.5	210 m - 470 m
21.0	170 m - 380 m	21.5	130 m - 300 m
22.0	110 m - 240 m	22.5	85 m - 190 m
23.0	65 m - 150 m	23.5	50 m - 120 m
24.0	40 m - 95 m	24.5	35 m - 75 m
25.0	25 m - 60 m	25.5	20 m - 50 m
26.0	17 m - 37 m	26.5	13 m - 30 m
27.0	11 m - 24 m	27.5	8 m - 19 m
28.0	7 m - 15 m	28.5	5 m - 12 m

Powell *et al.* (1997 [248]). NEOs as near Earth resources for mining are discussed by Gertsch *et al.* (1997 [108]).

10.2.4 The Cretaceous-Tertiary Impact

65 million years ago at the Cretaceous-Tertiary boundary (K/T) an impact occurred. Mass extinctions of a broad spectrum of lifeforms (Raup and Sepkoski, 1988 [252]), a worldwide clay layer containing geochemical (Alvarez *et al.*, 1980 [6]), mineralogical (Bohor, 1990 [42]) and isotopic anomalies (McDougall, 1988 [217]) and tsunami deposits (Bourgeois *et al.*, 1988 [46]) point to a major event at that time.

The buried Chicxulub basin is the source crater about 300 km in diameter. It is believed that the Chicxulub crater would most likely be formed by a long-period comet composed primarily of nonsilicate materials (ice, hydrocarbons etc.) and subordinate amounts primitive chondritic material. The collision would have raised the energy equivalent to between 4×10^8 and 4×10^9 megatons of TNT. Studies of terrestrial impact rates suggest that such an event would have a mean production rate of $\sim 1.25 \times 10^{-9}\,\mathrm{yr}^{-1}$. This rate is considerably lower than that of the major mass extinctions over the last 250 million years ($\sim 5 \times 10^{-7}\,\mathrm{yr}^{-1}$). However, there is substantial evidence establishing the cause-link between the Chicxulub basin forming event and the K/T biological extinctions. The crater showed several rings (similar to the rings of the Mare Orientale on the Moon).

Let us consider the impact of a small asteroid. It fragments in the atmosphere thus the cross section for aerodynamic braking is greatly enhanced. Ground impact damage such as

- craters,
- earthquakes,
- tsunami

from a stoney asteroid is negligible if it is less than 200 m in diameter. Small and relatively frequent impactors such as Tunguska produce only air blast damage and leave no long term scars. Objects 2.5 times larger which hit every few thousand years cause coherent destruction over many thousand km of coast. Let us assume an asteroid > 200 m hits an ocean. A water wave generated by such an impactor has a long range because it is two-dimensional; its height falls off inversely with distance from the impact. When the wave strikes a continental shelf, its speed decreases and its height increases to produce tsunamis. Tsunamis produce most of the damage from asteroids between 200 m and 1 km. An impact anywhere in the Atlantic by an asteroid 400 m in diameter would devastate the coasts on both sides of the ocean by tsunamis over 100 m high. An asteroid 5 km in diameter hitting the mid Atlantic would produce tsunami that would inundate the entire East coast of the US to the Appalachian mountains (see the paper of Hills and Mader, 1997 [133]).

In Fig.10.1 the estimated frequency of impacts as a function of asteroid diameter is shown.

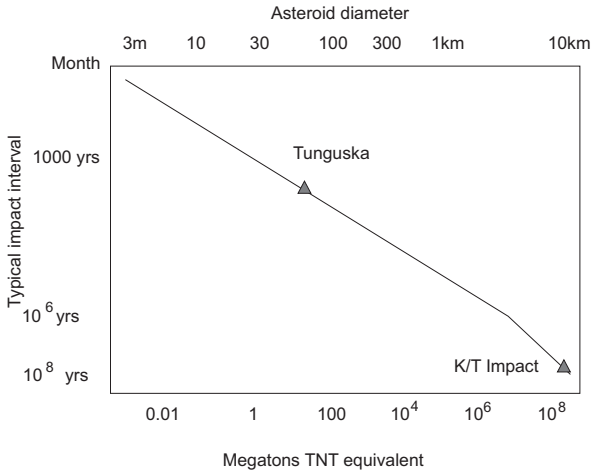


Figure 10.1: Estimated frequency of impacts on the Earth from the present population of comets and asteroids and impact craters.

An asteroid or comet ≥ 10 km in diameter (which releases $\geq 10^{24}$ J or 10^8 Mt TNT) would cause a global catastrophe:

- production of dense clouds of ejecta,
- smoke clouds,
- large amounts of nitric oxides are created $\rightarrow$ acid rain;
- the NO_x in the stratosphere destroys the ozone layer.
- in an ocean impact: enhanced greenhouse effect from water vapor injected into the atmosphere,
- CO_2 released by impact into carbonate rocks...

Some more details about such scenarios can be found e.g. in Chapman and Morrison (1994 [63]) or Melosh *et al.* (1990 [223]).

During the last 500 Million years there occurred several extinctions of marine species: in Table 10.3 we give the formal end of stage in Myr.

From the data we can deduce that the Earth should be hit by several asteroids and comets larger than a few km ($\sim 10^{23}$ J energy release) and perhaps one $\leq$10 km in a period of $\sim$ 100 Myr. Thus for the last 500 Myr years 5 events of extinctions are to be expected and 20 minor events which is in agreement with astronomical predictions.

From normal meteoroid ablation an iridium anomaly is observed and is one of the most significant signatures for impact. The search for iridium has resulted in reports of elevated iridium levels ($\leq$ 10 times background values) at or near a

Table 10.3: Extinction of marine species. The end of the stage is given in Myr.

Name	End	Name	End
Pliocene	2.3	Mid Miocene	11.2
Upper Eocene	34	Maastrichtian	65
Cenomanian	93.5	Aptian	112
Tithonian	144	Callovian	159
Pliensbachian	190	Norian	206
Carnian	221	Tatarian	248
Guadelupian	250	Stephanian	290
Serpukhovian	322	Famennian	363
Frasnian	367	Eifelian	380
Ludlovian	411	Ashgillian	439
Llanvirnian	469	Tremadocian	493
Trempeleauean	505	Franconian	508
Botomian	520	Proterozoic/Cambrian	540

number of extinction boundaries. The Ir levels are generally significantly weaker than the K/T anomaly.

The conversion from projectile mass to crater size was given by Shoemaker *et al.* (1990 [284]):

$$D = 0.074 C_f (g_e/g)^{1/6} (W \rho_a/\rho_t)^{1/3.4} \tag{10.1}$$

D...crater diameter, c_f...crater collapse factor (1.3 for craters larger than 4 km on Earth), g_e...gravitational acceleration at the surface of the Earth, g... acceleration at the surface of the body on which the crater is formed (in this case Earth), W... the kinetic energy of the impacting body in kilotons TNT, ρ_a...density of the impactor ($1.8\,\mathrm{g/cm^3}$ for a comet...$7.3\,\mathrm{g/cm^3}$ for an iron body) and ρ_t...density of the target rock (e.g. $2.7\,\mathrm{g/cm^3}$). From this equation we see, that a carbonaceous chondrite would produce a crater $\sim$ 94 km in diameter whereas an impactor of cometary composition $\sim$ 150 km.

10.3 Meteorites

10.3.1 General Properties

The term meteor comes from the Greek meteoron, meaning phenomenon in the sky. Meteors are small solid particles that enter the Earth's atmosphere from interplanetary space. They move at high speeds and the friction they encounter in the air vaporizes them (typically at heights between 80 and 110 km above the surface). The light caused by the luminous vapors formed in such an encounter appears like a star moving rapidly across the sky, fading within a few seconds. To be visible, a meteor must be within 200 km of the observer. The total number of meteors bright enough to be visible is estimated to be about 25 million per day.

A meteoroid is matter revolving around the sun or any object in interplanetary space that is too small to be called an asteroid or a comet. Even smaller particles are called micrometeoroids or cosmic dust grains, which includes any interstellar material that should happen to enter our solar system. A meteorite is a meteoroid that reaches the surface of the Earth without being completely vaporized.

One of the primary goals of studying meteorites is to determine the history and origin of their parent bodies. Several achondrites sampled from Antarctica since 1981 have conclusively been shown to have originated from the moon based on compositional matches of lunar rocks obtained by the Apollo missions of 1969-1972. Sources of other specific metorites remain unproven, although another set of eight achondrites are suspected to have come from Mars. These meteorites contain atmospheric gases trapped in shock melted minerals which match the composition of the Martian atmosphere as measured by the Viking landers in 1976. All other groups are presumed to have originated on asteroids or comets; the majority of meteorites are believed to be fragments of asteroids.

A typical bright meteor is produced by a particle with a mass less than 1 g. A particle the size of a golf ball produces a bright fireball. The total mass of meteoritic material entering the Earth's atmosphere is estimated to be about 100 tons per day.

10.3.2 Classification

Meteorites can be classified into stony, stony iron and iron. The most common meteorites are chondrites which are stony. Radiometric dating indicate an age of about 4.5×10^9 years. Achondrites are also stony but they are considered differentiated or reprocessed matter. They are formed by melting and recristallization on or within meteorite parent bodies. Pallasites are stony iron meteorites composed of olivine enclosed in metal.

The motion of meteoroids can be severely perturbed by the gravitational fields of major planets. Jupiter's gravitational influence is capable of reshaping an asteroid's orbit from the main belt so that it dives into the inner solar system and crosses the orbit of Earth. This is apparently the case of the Apollo and Vesta asteroid fragments.

Particles found in highly correlated orbits are called stream components and those found in random orbits are called sporadic components. It is thought that most meteor streams are formed by the decay of a comet nucleus and consequently are spread around the original orbit of the comet. When Earth's orbit intersects a meteor stream, the meteor rate is increased and a meteor shower results. Typically, a meteor shower will be active for several days. A particularly intense meteor shower is called a meteor storm. Sporadic meteors are believed to have had a gradual loss of orbital coherence with a meteor shower due to collisions and radiative effects, further enhanced by gravitational influences. There is still some debate concerning sporadic meteors and their relationship with showers. A well known meteor shower are the Perseids (named after the radiant that is the name of the constellation where the meteorites are seem to be coming from) which has its maximum on August 11.

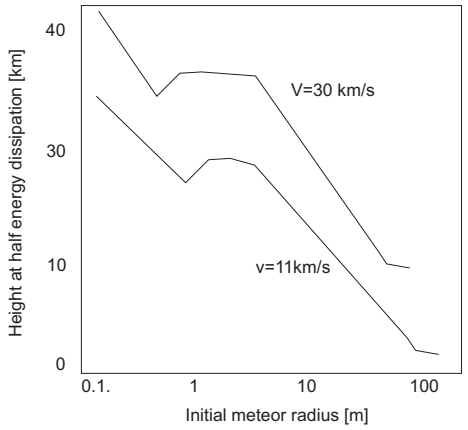


Figure 10.2: Height in the atmosphere at which half the kinetic energy of a stony meteoroid is dissipated. Note that asteroids with > 100 m hit the ground with most of their original kinetic energy.

10.3.3 The Leonid Threat

On Nov. 18th the Earth glides into a dust cloud shed by the comet Tempel-Tuttle in 1766. Around this date lots of meteors can be seen. The Leonid meteor shower is seen to emanate from a point in the western part of Leo which is called the radiant. When Leonids rain down on the airless moon, they will not cause the shooting stars because there is no atmosphere; they just hit the ground. In 1999 while the Moon passed through the Leonid debris, impact flashes were recorded. When such a meteoroid hits the Moon it vaporizes some dust and rock. Some of those vapors contain sodium (constituent of lunar rocks) which scatters sunlight. The Moon is also surrounded by a gaseous halo called the lunar exosphere ($\sim$ 100 atoms per cm^3). The solar wind blows it into a long tail (much like a comet tail). This tail points away from the Sun and extends several 10^5 km and the Earth passes through it once a month around the time of New Moon. Using extraordinary sensitive cameras, sunlight scattered can be detected. After the Leonid fireball shower of 1998 the density of the Moon's sodium trail tripled.

There is also a risk to satellites. The damage can be:

- mechanical: direct impact through spalling or chipping as larger particles hit the spacecraft and break up;
- electrical: electrostatic discharges (ESDs), electromagnetic pulse (EMP).

Mechanical damage consists predominantly of sandblasting which all spacecraft experience during the Leonids. This causes surface degradations. This is in general not a serious problem. Impacts and spall result from larger particles hitting a satellite. These can punch holes in a solar panel or wall. Spall produced by secondary particles can affect the internal mechanisms of the spacecraft more seriously than the original impact.

EMPs are created from the direct vaporization of impacting particles into plasma and ESDs by a buildup of charging over the satellite surface. Both can cause electrical and communication problems, erroneous signals in telemetry and short circuits.

How can such problems be minimized? One simple manoeuvre is to minimize the cross-sectional area of the satellite that is exposed to the meteor shower; e.g. the solar panels point edge on into the meteor stream by reorienting the spacecraft. Another technique is to turn off equipment that is particularly sensitive to ESDs.

Meteoroid impacts on spacecrafts and penetration damage are studied by McBride and McDonnell (1999 [214]). Meteoroid morphology and density (e.g. using NASA's LDEF satellite results) were investigated by McDonnell and Gardner (1998 [216]). The Leonid Meteor Shower and the Lunar Sodium Atmosphere are treated in Hunten *al.*, 1998 [143]. Charge production, impact probability, the impact penetration depth, the amount of ionization generated upon impact on spacecraft due to the strong Leonid shower in 1999 and 2000 are calculated by McNeil *et al.*, 1998 [221]

10.4 Comets

10.4.1 General Properties

Bright comets may be easily seen with the naked eye[4]. Comets consist of different parts:

- nucleus: diameter between 1 and 50 km
- Coma: size in the range of 10^4 – 10^5 km.
- Hydrogen coma: not seen with naked eye, surrounds the visible gas/dust coma, several million km.
- Tail: there are two different types of tails:
  1. yellowish dust tail with an extension 10^6–10^7 km
  2. bluish plasma tail, or ion tail, extension 10^7–10^8 km

Comets are named after their discoverers and are also given a number, the year of discovery (or recovery) and a letter. In this designation 2006c denotes the third comet discovered in 2006. There is another designation where comets are distinguished by their perihelion passage: 2006X is the 10th comet in 2006 according to perihelion passage. Periodic comets are by a P/ and comets which have collided with other bodies by D/.

[4]see also: Comets and the Origin and Evolution of Life, by P. J. Thomas, Springer, 2006

10.4.2 Cometary Activity

Cometary activity is triggered by the Sun. At large heliocentric distances comets cannot be discerned from asteroids. The cometary nucleus is covered with ice which sublimates[5] and dust is dragged on. Thus the coma is formed. The heliocentric distance where this process occurs varies from comet to comet. Many comets start to become active when passing the orbit of Jupiter, some comets were observed with a coma at a distance beyond Uranus. As a rule of thumb, most comets become bright when getting closer than 3 AU. The brightness of a comet obeys the law:

$$m \sim \frac{1}{r_\odot^\zeta r_\Delta^2} \tag{10.2}$$

In this equation r_Δ is the distance to the observer and $r_\odot$ the heliocentric distance. Note, that the dependence is $\sim 1/r_\odot^\zeta$ with $\zeta > 2$. This takes into account, that the gas production increases strongly with decreasing distance from the sun.

Cometary tails point in the direction opposite to the Sun. Radiation pressure from the Sun acts on molecules via absorption and re emission of solar photons. The dust particles orbit independently around the Sun, the attraction to the Sun is diminished by the radiation pressure. Therefore, dust tails are broad and curved.

The ion tail of comets is caused by the solar wind. By solar UV photons, neutral atoms are ionized e.g.

$$\mathrm{CO} + \mathrm{h}\nu \rightarrow \mathrm{CO}^+ \tag{10.3}$$

Since magnetic fields are carried by the solar wind, these ions are susceptible to the magnetic force. Because the most common ion CO^+ scatters blue light better than red, the ion tail appears bluish. The solar wind sweeps past comets at 500 km/s

10.4.3 Oort Cloud and Kuiper Belt

Oort proposed in 1950, that comets reside in a vast cloud at the outer reaches of the solar system. This has come to be known as the Oort Cloud. This hypothesis is based on several observational facts: a) no comet has been observed with a hyperbolic orbit (which would indicate interstellar origin), b) aphelia of long period comets lie at a distance of about 50 000 AU, c) there is no preferential direction from which comets come. The Oort cloud may contain up to 10^{12} comets (in total about the mass of Jupiter). The dynamical lifetime of an object in this cloud to ejection by passing stars can be estimated to be half of the age of the solar system. For such estimation also encounters with molecular clouds have to be considered. The objects in the Oort cloud need to be replenished either by capture from the interstellar medium or by a vast inner cloud. It is assumed that encounters with dense interstellar clouds could cause perturbations of the inner cloud and replenish the outer cloud. During such perturbations "showers" of comets penetrate to the inner solar system. This could happen every 10^8 years.

[5] sublimation means direct evaporation from the solid state

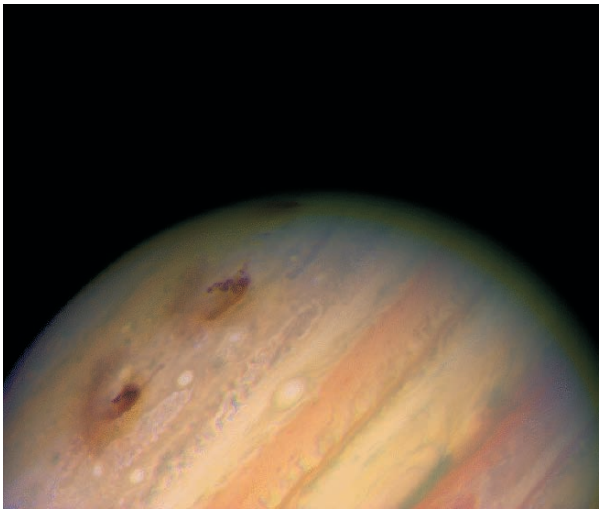


Figure 10.3: Impact of comet Shoemaker-Levy fragment on Jupiter. NASA/HST

Comets from the Oort cloud may pass up to 400 times the inner solar system before the object becomes perturbed by the inner planets and is changed into a short period comet.

From the observed short period comets the total mass of bodies in the Kuiper belt is estimated to be 0.0026 Earth masses which corresponds to 10^9 to 10^{10} objects. The Kuiper Belt is a disk-shaped region past the orbit of Neptune roughly 30 to 100 AU from the Sun containing many small icy bodies.

10.4.4 Comets and Meteor Showers

Comets dissolve leaving a cloud of debris behind their orbit. When the Earth crosses cometary orbits, meteor showers occur. The most famous is the Perseid shower. The shower begins, in mid-July when Earth enters the outskirts of a cloud of debris from Comet Swift-Tuttle. Dust-sized meteoroids hitting the atmosphere will streak across the night sky, at first only a few each night, but the rate will build. By August 12th when the shower peaks, sky watchers can expect to see dozens, sometimes even hundreds, of meteors per hour. The idea that comets and asteroids might threaten our planet was not widely accepted until the 1980s. Comet Swift-Tuttle is big, about the same size as the asteroid that wiped out dinosaurs 65 million years ago, and as recently as 1992 it seemed that Swift-Tuttle might strike Earth in the year 2126. New data and calculations show otherwise, though. There's no danger of a collision for at least a millennium and probably much longer. In 1994, July 16-22, over 20 fragments of the comet Shoemaker-Levy collided with Jupiter. This was observed worldwide (see Fig. 10.3).

The solar heliospheric observatory SOHO satellite observed many collisions of comets with the Sun (Fig. 10.4).

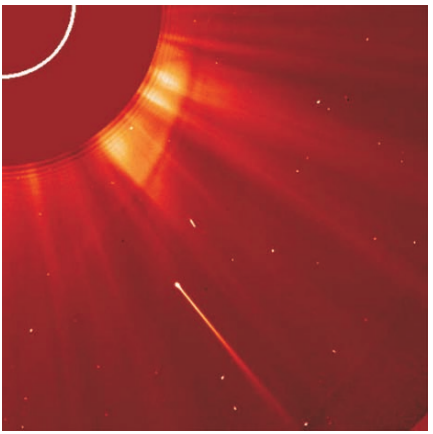


Figure 10.4: Comet observed by SOHO colliding with the Sun. SOHO/ESA, NASA

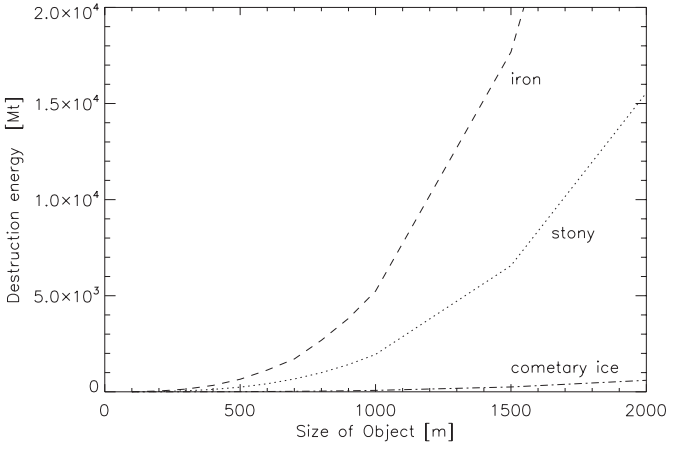


Figure 10.5: Destruction energies required for different bodies

Sazonov and Yakovlev (2006 [270]) discuss comets on dangerous orbits and how to change their orbits by explosive and sublimation methods. It is clear that for large bodies the required charge power to destroy the dangerous bodies is too large, and therefore, by means of an explosive impulse, such bodies can be moved to a safe trajectory. The required destruction energy depends on

- diameter of the object
- compositions of the object. One can calculate this for different objects such as stony, iron and cometary ice composition.

The energies are plotted in Fig. 10.5. For comparison: the energy of the Hiroshima bombe was 15 kt. A cometary composition body of 100 m diameter requires 0.075 Mt for destruction. 1 kt TNT is the equivalent to $4{,}184 \times 10^{12}$ J. The sum of all bombs during the second world war can be estimated to 2 Mt. The biggest H-bomb had 50 Mt.

Chapter 11

Space Debris

On Dec 3, 2001 BBC reports, that space debris lit up the sky. The spectacular nighttime light show seen over parts of southern England is now believed to have been caused by burning Russian space debris. Observers said the fragments, which could be seen over parts of Essex and Sussex, were very bright and traced across the sky for up to four minutes.

Orbital debris is defined as any man-made object in orbit around the Earth which no longer serves as a useful purpose.

We will discuss estimations on the number of debris elements as well as models that calculate their orbits and shielding mechanisms for spacecraft[1].

11.1 Number of Space Debris

11.1.1 Orbits

There are several types of satellite orbits serving different purposes. The inclination of an orbit is defined as: 0 degrees means an equatorial orbit, 90 degrees a polar orbit.

Polar orbits: these orbits allow the satellite to observe nearly every part on the Earth. The Earth rotates under the satellite. The inclination of the satellite is nearly 90 degrees. One orbit around the Earth is completed in approximately 90 minutes.

Sun synchronous orbits: a satellite will pass over a section of the Earth at the same time a day at a height between 700 and 800 km. Due to the revolution of the Earth around the Sun, the satellite has to shift its orbit approximately 1 degree per day- additional gravitational forces due to the bulge of the Earth at the equator are used for that acceleration.

Geosynchronous orbits: at a distance of 35790 km above Earth the satellites circle the Earth at the same rate as the Earth spins (23 hours, 56 minutes, and 4.09 seconds). These satellites observe almost a full hemisphere of the Earth,

[1] see e.g. the book: Space Debris Models and Risk Analysis, by H. Klinkrad, Springer 2006.

Figure 11.1: Orbits of GPS satellites. The satellites orbit the earth with a speed of 3.9 km/s. One revolution takes 12 h sidereal time, corresponding to 11 h 58 min earth time. This means that the same satellite reaches a certain position about 4 minutes earlier each day. The mean distance from the middle of the earth is 26560 km. The system consists of at least 24 satellites, the first one started in 1978. The European GALILEO will be operational in 2010, consisting of 30 satellites (27 working, 3 in reserve). *http : //www.kowoma.de/en/gps/orbits.htm*

are used to study large scale phenomena such as hurricanes, or cyclones and for communication satellites. The disadvantage of this type of orbit is that the Earth can be observed from there with low resolution.

On the other hand, satellites in Low Earth Orbit, LEO, can only cover a small area with high resolution and therefore often constellations are used, such as the GPS, IRIDIUM (66 active communication satellites, first launched in 1998). The current positions over 900 satellites can be followed using an online tool of NASA[2].

11.1.2 Number of Objects

In 1957 Sputnik 1 was launched as the first man made spacecraft. In the years of space activities some 3 750 launches led to more than 23 000 observable space objects (larger than 10 cm) of which currently 7 500 are still in orbit. Only 6% of the catalogued orbit population comprise operational spacecraft, while 50% can be attributed to decommissioned satellites, spent upper stages, and mission related objects (launch adapters, lens covers, etc.). The remainder of 44% is originating from 129 on-orbit fragmentation which have been recorded since 1961. These events, all but 1 or 2 of them explosions of spacecraft and upper stages, are assumed to have generated a population of objects larger than 1 cm on the order of 70 000 to 120 000. Only in the range of 0.1 mm size the sporadic flux from meteoroids prevails over man-made debris. From a statistical point of view we have to note

[2] *http : //science.nasa.gov/Realtime/jtrack/3d/JTrack3D.html*

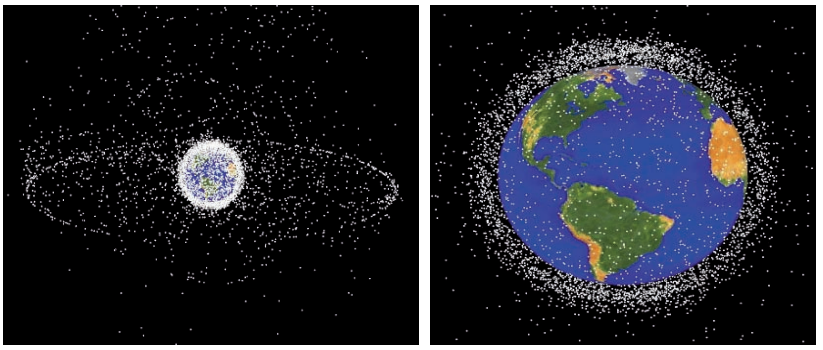


Figure 11.2: GEO and LEO objects as a source of space debris. GEO denotes geostationary orbit.

that most orbital debris reside within 2 000 km of the Earth's surface. Within this volume, the amount of debris varies significantly with altitude and regions of debris concentration are found near 800 km, 1 000 km and 1 500 km.

From the above considerations it is clear that spacecrafts have to be protected from collisions with space debris. Let us mention two examples: the US space command examines the trajectories of the Space Shuttle in order to identify possible close encounters with space debris. If a dangerous object is believed to approach a few tens of kilometers to the Space Shuttle, it will be maneuvered away from the object (although in such a case the chances of a collision are only approximately 1:100 000). Such an operation is necessary about once every year or two (at present).

Space debris is an inherently international problem and its solution requires international co-operation. The Inter-Agency Space Debris Coordination Committee (IADC) whose members are ESA, NASDA (Japan), NASA, and the Russian Space Agency RKA and the Canadian Space Agency (CSA) provides a forum for discussion and coordination of technical space debris issues.

11.2 Detection of Space Debris

Remote sensing of space debris from ground-based measurements falls into two categories:

- Radar measurements: these have been used for space debris in low Earth orbit (LEO).
- Optical measurements: these have been used for high Earth orbit (HEO). For passive optical measurements the intensity of the signal from space debris is inversely proportional to the square of its distance or altitude:

$$I_{\text{optical}} \sim 1/r^2 \tag{11.1}$$

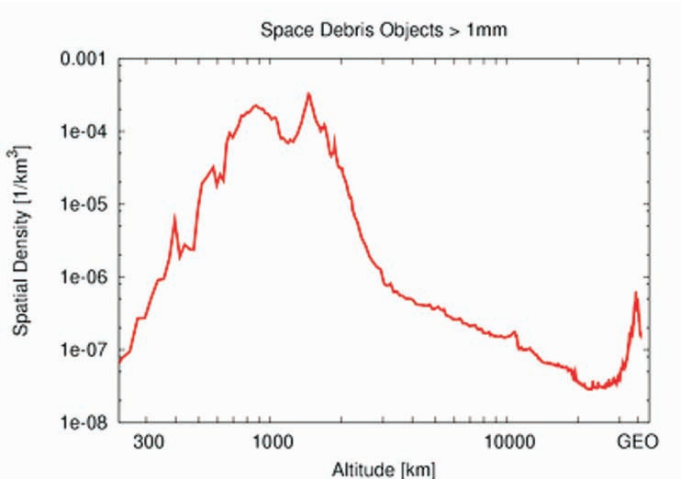


Figure 11.3: Spatial density of space debris by altitude according to ESA MASTER-2001.

The incident illumination from the Sun is essentially independent of altitude. For radar measurements:

$$I_{\text{radar}} \sim 1/r^4 \tag{11.2}$$

since radars must provide their own illumination.

Therefore, optical telescopes of modest size are more suitable than most radars for detection of debris at high altitudes. On the other hand, radars are better suited to detect objects in LEO.

11.2.1 Radar Measurements

Ground-based radars are well suited to observe space objects because

- all weather,
- all day-and-night

performance. There are two types for space object measurements:

1. Radars with mechanically controlled beam direction using parabolic reflector antennas; here, only objects in the field of view (which is given by the mechanical direction of the parabolic reflector antenna) can be observed; used for tracking or imaging satellites.

2. Radars with electronically controlled beam direction using phased array antennas. In that case multiple objects at different directions can be detected and measured simultaneously; used for tracking and search tasks.

In the tracking mode the radar follows an object for a few minutes gaining data on angular direction, range, range rate, amplitude and phase of the radar echoes. From these parameters the orbital elements can be derived In the beam-park mode, the antenna is kept fixed in a given direction and echoes are received from objects passing within its field of view. This yields statistical information on the number and size of detected objects; the determination of the orbit is less precise. There is also a mixed mode. From the radar measurements the following parameters can be derived:

1. orbital elements; thus the motion of the object's center of mass around Earth is defined.

2. Attitude; describes the motion of the object around its center of mass.

3. Size and Shape of the object.

4. Ballistic coefficient; this describes the rate at which the orbital semi-major axis decays.

5. Object mass,

6. material properties.

The main source of data for space debris in the size range of 1-30 cm is the NASA Haystake radar facility operated by MIT Lincoln Laboratory. Under an agreement with the US Air Force since 1990 data are collected. The data indicate that there are about 100 000 fragments in orbits with sizes down to 1 cm.

When space debris (man made or meteoroids) enter the atmosphere an ionization trail is created. Molecules in the upper atmosphere are ionized by the passage of the meteor. Such ionization trails can last up to 45 minutes. Such an ion trail will act as a mirror for radio waves. Radar measurements operating at 50 MHz permit to estimate the reentry of space debris, the ionization trail behind reentering bodies can be detected. With such facilities one can detect meteors as small as 100 microns.

Radar measurements of space debris have been done at Haystack (US) and Goldstone radars (US), Russia and by Germany using the Research Establishment for Applied Science (FGAN) radar and the Effelsberg radio telescope. Haystack and Goldstone radars have provided a statistical picture of LEO debris at sizes down to 0.5 cm which was confirmed by FGAN. These measurements have proven that the debris population exceeds the natural meteoroid population for all sizes (except between 30 and 500 m).

Radar measurements and the usage of the Sardinian Radio Telescope for space debris detection are described e.g. in the article of Di Martino *et al.*, 2006 [74].

The international radar space debris research was reviewed by Molotov *et al.* (2005 [224]) where further literature is cited.

11.2.2 Telescopes

Space debris can be categorized into objects that reflect radar well but sunlight poorly. The other group reflects sunlight well but radar poorly. Thus, radar and optical telescopes see somewhat different debris populations. With the use of optical telescopes, debris at very high altitudes (e.g. in geosynchronous orbits, GEO) can be detected.

The US Space Command employs aperture telescopes of 1 m to track HEO objects. With these telescopes objects of 1 m at geosynchronous altitudes, corresponding to a limiting stellar magnitude of 16 can be detected. A limiting stellar magnitude of 17 or greater is needed to detect debris smaller than 1 m near GEO.

Most objects in GEO are intact; in 1978 a Russian Ekran satellite in GEO was observed to explode.

NASA is using two optical telescopes for measuring orbital debris[3]: a 3 m diameter liquid mirror telescope which is referred to as the LMT, and a charged coupled device-equipped 0.3 m Schmidt camera, which is commonly referred to as the CCD Debris Telescope or CDT. The LMT consists of a 3 m diameter parabolic dish that holds 14 l of liquid mercury. The dish is spun up to a rate of 10 revolutions per minute. Centrifugal force and gravity cause the mercury to spread out in a thin layer over the dish creating a reflective parabolic surface that is as good as many polished glass mirrors.

11.2.3 Catalogues

There are two catalogues of space objects that are frequently updated:

- United States Space Command catalogue,
- Space Object catalogue of the Russian Federation.

Based on those two catalogues data are also archived in the Database and Information System Characterizing Objects in Space (DISCOS) of ESA. The National Space Development Agency (NASDA) of Japan is studying a debris database. Current catalogues contain information on satellites and debris as small as 10-30 cm in diameter. Some recent activities are aimed to provide detection of 5 cm objects at altitudes below 600 km. For smaller sizes modelers must use statistical measurements.

11.3 Shielding and Risk Assessments

11.3.1 Risk Assessments

Risk assessments are utilized in the design of manned and unmanned spacecraft. They aid in the placement and protective shielding design. This is of course only feasible for critical subsystems and components. It becomes extremely important in the system design of large communication satellite constellations. In Table 11.1 a summary of the studies made so far is given.

[3]NASA Orbital Debris Observatory (NODO)

Table 11.1: Mean time between impacts on a satellite with a cross-section area of $10\,m^2$

Height of circular orbit	Objects 0.1-1.0 cm	Objects 1-10 cm	Objects >10 cm
500 km	10-100 yrs	3 500-7 000 yrs	150 000 yrs
1 000 km	3-30 yrs	700 - 1 400 yrs	20 000 yrs
1 500 km	7-70 yrs	1 000-2 000 yrs	30 000 yrs

For GEO the situation is more complicated. The number of space debris of less than 1 m in diameter is not well known. Moreover, there is no natural removal mechanism for satellites in GEO. One can estimate an annual collision probability for an average operational satellite with other catalogued objects at 10^{-5}.

Another problem concerns the re-entry. Since the last 40 years 16 000 re-entries of catalogued space objects are recorded. No significant damage or injury occurred which can be attributed to the large expanse of ocean surface and sparse population density in many land regions. During the past years, approximately once each week an object with a cross section of $1\,m^2$ or more entered the Earth's atmosphere. The risk of re-entry comes from:

- Mechanical impact,
- chemical contamination,
- radiological contamination.

Since about 12% of the present catalogued space debris population consists of objects discarded during normal satellite deployment (fasteners, yaw, weights, nozzle covers, lens caps, tethers,...) one should take mitigation measures against these objects. 85% of all space debris larger than 5 cm result from fragmentation of upper stages. In 1996 the French CERISE spacecraft was struck and partially disabled by an impact fragment which most probably came from an exploded Ariane upper stage.

A family of space debris objects was found at a height of 900 km. The density peak found there is caused by a large number of sodium-potassium liquid metals droplets- they have been used as a coolant for the on board nuclear reactor, leaked from the Russian ocean surveillance satellites. The estimation is about 70000 drops with diameters between 0.5 mm and 5.5 and the detection was mainly made with the Haystack radar. Using the Goldstone radar the so called West Ford Needles at an altitude of 2900 km were detected. They are copper dipoles, 1.77 cm long, and remnants that were released in 1961 and 1963 by the US MIDAS 3 and MIDAS 6 satellites for telecommunication experiments. It was first expected that they should reenter the Earth's atmosphere within 5 years but now a population of 40000 objects were found between 2400 and 3100 km.

In total it is estimated that more than 350000 objects larger than 1 mm crowd the space around Earth and a particle around 5 mm is able to directly penetrate the shuttle cabin. For more details see Valsecchi *et al.*, 2006 [323].

Collision risk assessment for a spacecraft in space debris environment was studied e.g. by Tang *et al.*, 2005 [311]

11.3.2 Reentry of Orbital Debris

How long will orbital debris remain in Earth orbit? As a rule of thumb one can say that the higher the altitude, the longer the debris will typically remain in Earth orbit.

- Debris left in orbits below 600 km: normally falls back to Earth within a few years.
- Debris left in orbits at altitudes of 800 km: the time for orbital decay is several decades.
- Debris left in orbits at altitudes above 1 000 km: will normally continue circling the Earth for a century or more.

Up to now no serious injury or property damage has been confirmed caused by reentering debris. Most of the space debris does not survive the severe heating which occurs during reentry. During the past 40 years, on the average one cataloged piece of debris fell back to earth each day.

On 12 June 1979 Skylab (70 t) came crashing to Earth, scattering chunks of metal over the West Australian desert. US officials were unable to control it's final descent. Pieces of the Russian space station Mir could be observed racing across the sky above Fiji as Mir made its descent into the earth's atmosphere on March 23, 2001. Mir plunged to earth after Russian Mission Control fired engines to nudge it out of the orbit it had kept for 15 years. The entrance velocity was 6 400 km/h and the final burst of rockets was made at a height of only 170 km over Africa. The weight of the space station was about 135 tons.

A very spectacular event was the crashing of the Russian satellite KOSMOS 954, an active radar satellite for ocean surveillance. The high power consumption of the active radar required a nuclear reactor as power source. The reactor of KOSMOS 954 reentered over a desert area in Northern Canada and an area of 124 000 m^2 was contaminated.

The Soviet Mars 96 satellite that had 270 g of the poisonous plutonium on board crashed in march 1996 at a distance of 1300 km west of South America into the Pacific.

A survey of some events is given in Table 11.2.

The crashes of EUVE and Compton were controlled by NASA.

11.3.3 Orbital Debris Protection

Many efforts are made to develop protection:

- hypervelocity impact measurements: in such experiments projectiles are produced at speeds more than seven times faster than the fastest bullet; this is done with so called two stage light gas guns. The impact event lasts only

Table 11.2: Spectacular satellite crashes on Earth.

Name	Crash in	Location of Crash	Reported damage
KOSMOS 954	1978	North Canada	radioactive
Skylab	1979	Australia	Cow was killed
Chinese Satellite	1996	Atlantic	
Mars 96	1996	Pacific	Plutonium
Mir	2001	Pacific	
Chinese Satellite	2004	China	house damaged
Compton γ Ray Obs.	2000	Pacific	15 t
EUVE	2002	South Pacific	3 t

a few microseconds. The velocity of the bullet is measured by using two laser curtains positioned a short distance uprange of the target. The distance between the curtains is known and the time elapsed between the two disruptions is measured, thus the projectile velocity can be measured.

- Shield development.

- Simulations: sophisticated computer programs simulating hypervelocity events are run on supercomputers. This approach to developing spacecraft shield solutions is becoming more and more prevalent.

- Developing new materials

- Impacts on spacecraft: all spacecraft collide with very small orbital debris particles and meteoroids. The Long Duration Exposure Facility (LDEF) was a bus sized spacecraft. It was returned after 5.7 years in low Earth orbit. The LDEF was placed in low Earth orbit (LEO) by the space shuttle Challenger in April 1984 and retrieved by the space shuttle Columbia in January 1990. On the LDEF over 30 000 impacts were found (these craters were visible to the naked eye and larger than 0.5 mm). Form that sample about 1000 were chemically analyzed in order to investigate the origin of the projectiles. The largest crater found on LDEF had a diameter of 5 mm and was probably caused by a particle of 1 mm. Some impacts were clustered in time. On the European Retrievable Carrier (EURECA), the largest impact crater diameter was 6.4 mm. The returned solar array of the HST (Hubble Space Telescope, NASA/ESA) had been the one with the highest orbit altitude. It was found that the impact flux for HST was considerably higher (factor 2-8) than for EURECA. The infra-red astronomical satellite (IRAS), launched in 1983 to perform a sky survey at wavelengths ranging from 8 to 120 μm was operational during 10 months near altitude of 900 km. 200 000 potential debris sightings are stored in a database. About 10 000 sightings are attributed to real objects. A plot of debris flux in low Earth orbit as a function of object size (cm) is given in Fig. 11.4 where the coordinates are logarithmic.

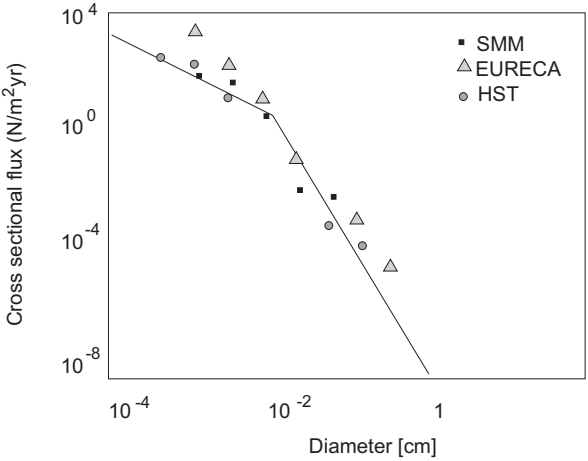


Figure 11.4: Approximate measured debris flux in low Earth orbit by object size (sketch)

Table 11.3: Some examples of retrieved spacecraft and surfaces

Name	Orbit	In orbit	Exposed area
Salyut 4,6	350 km	1974-1979	$\sim 7\,\mathrm{m}^2$
STS-7 Window (NASA)	295-320 km	June 1983	$\sim 2.5\,\mathrm{m}^2$
SMM (NASA)	500-570 km	1980-1994	$2.3\,\mathrm{m}^2$
LDEF (NASA)	340-470 km	1984-1990	$151\,\mathrm{m}^2$
EURECA (ESA)	520 km	1992-1993	$35\,\mathrm{m}^2$
HST (solar array)	610 km	1990-1993	$62\,\mathrm{m}^2$
Mir	390 km	1986-1998	$\sim 15\,\mathrm{m}^2$

- analysis of returned spacecraft surface; Critical surfaces, such as the windows, on the Space Shuttle are examined after every flight.

  Donald H. Humes and William H. Kinard from NASA Langley Research Center examined the WF/PC-I radiator with a microscope to measure the damage done by meteoroids and man-made orbital debris during its 3.6 years in orbit. They measured about 100 possible impact sites and rated them by size on an arbitrary scale of 1 to 10 (10 being the largest). They found 14 impact craters with a diameter greater than 450 microns.

At NASA a hypervelocity impact technology facility is under operation (HITF).

Of course the International Space Station (ISS) will be the most heavily shielded spacecraft ever flown. Critical components (e.g., habitable compartments and high press tanks) will normally be able to withstand the impact of debris as large as 1 cm in diameter. ISS will also have manoeuvering capability to avoid hazardous objects.

11.3.4 Space Debris Models

To assess the risk potential of collisions of man-made or natural particulates with operational spacecraft, one must refer to statistical models of the particle population for all size regimes except for man-made debris above 10 cm. In the latter case, collision events or near-miss events can be predicted on the basis of orbital data from operational surveillance networks of. In the former case, collision fluxes can only be estimated statistically. Currently, space debris between 1 cm and 10 cm are neither observable, nor are they shieldable with available on-orbit technology. Hyper-Velocity Impact (HVI) tests are used to experimentally verify and improve shields for on-orbit use, with the aim to increase the shieldable impactor size beyond 1 cm. The Mission Analysis Section of ESOC is coordinating all Space Debris Research Activities within ESA. SOC's Meteoroid and Space Debris Terrestrial Environment Reference (MASTER) model can be used to assess the debris or meteoroid flux imparted on a spacecraft on an arbitrary earth orbit.

At NASA, a new modelling technique called Smooth Particle Hydrodynamics (SPH) is under development (Hyde and Christiansen, 2002 [145]). Their approach models the distribution of debris fragments from a collision without using the normal computational mesh that is often subject to tangling. SPH eliminates many difficulties of previous calculation techniques.

The four main activities of ESA-ESOC space debris task group are:

- development of a meteoroid and debris reference model;
- radar measurements of mid-size debris; these are necessary since current models in low earth orbit suffer from significant uncertainties about objects smaller than about 50 cm. This is essential for spacecraft which require protection; it is currently technically not feasible to shield against objects larger than 1 cm. The feasibility of detecting and tracking medium-size debris (1 to 50 cm)with a high power radar at the Forschungsgesellschaft für Angewandte Naturwissenschaften (FGAN)in Germany was investigated.
- Optical measurements; these are suited for objects in high altitude orbits. The detectors use CCD and a 1 m telescope will be operated by ESA at the Teide observatory in Tenerife.
- Analysis of spacecraft surfaces returned from space

The main aim of the mathematical model is a description of the debris and meteoroid environment at altitudes between low Earth orbit (LEO) and the geostationary orbit (GEO). The minimum size of an object is 0.1 mm. The model is based on the catalogued population and on known break-ups of spacecraft and rocket upper-stages in orbit. The initial distribution of fragments is described in terms of their position, velocity, mass. The objects are then propagated forward in time taking into account the relevant perturbations. The description of MASTER (Meteoroid and Space Debris Terrestrial Reference Model) is given in Sdunnus *et al.* 2001 [276].

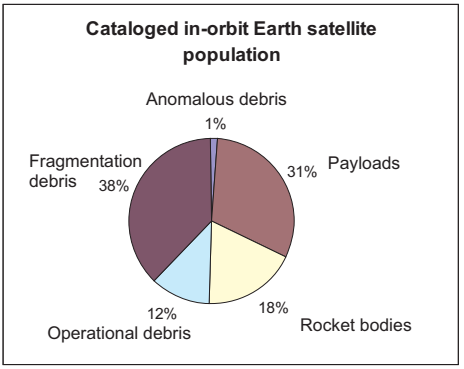


Figure 11.5: Segments of the cataloged in-orbit Earth satellite population.

11.3.5 Shielding

Protection against particles 0.1-1 cm size can be achieved by shielding spacecraft structures. Objects 1-10 cm in size cannot be shielded nor can they be routinely tracked by surveillance networks. Protection against these particles can be achieved through special features in the design (e.g. redundant systems, frangible structures...). Physical protection against particles larger 10 cm is not technically feasible. In front of the spacecraft wall single sheet Whipple bumbers or complex layers of metal and ceramic/polymer fabrics can be used for shieldings. They break up the impacting particle and absorb the energy of the resulting ejecta. Bumper shields should be positioned at a sufficient distance from the shielded object.

The penetration depth (damage potential) of an impacting object depends on:

- mass,
- velocity,
- shape of the object; and of course
- material properties of the shield.

For manned spacecraft shield designs offer protection against objects smaller than 1 cm. The PNP (probability of penetration) is an important criterion for shield design. One can also install automatic detection systems to locate damage. For EVA (extravehicular activities) current spacesuits have many features with inherent shielding qualities to offer protection from objects of sizes up to 0.1 mm. By properly orientating their spacecraft, astronauts may also be able to use their vehicles against the majority of space debris or direct meteoroid streams. The United States Space Surveillance Network (SSN) and the Russian Space Surveillance System(SSS) monitor the LEO environment to warn crewed spacecraft if an object is projected to approach within a few km. If an object is predicted to pass through a box of $5 \times 225 \times 5$ km oriented along the flight path of the United States

Space Shuttle, the SSN sensor intensifies its tracking of the potential risk object. If the improved fly-by prediction indicates a conjunction within a box of $2 \times 5 \times 2$ km an avoidance manoeuvre is performed. During 1986-1997 4 such evasive manoeuvres were executed. Collision avoidance manoeuvres were performed by the ESA satellite ERS-1 in June 1997 and March 1998 and by the CNES satellite SPOT-2 in July 1997.

Calculations made prior to the launch of spacecrafts permit the establishment of safe launch windows.

For unmanned spacecraft, lower PNPs are tolerable.

The necessity for collision avoidance manoeuvres was already pointed out by Rex *et al.*, 1991 [255].

An overview of fragmentation of LEO Upper Stages was given by Chernyavskiy *et al.* (1994 [64]).

A technical report on space debris was given from the Scientific and Technical Subcommittee of the United Nations Committee on the Peaceful uses of Outer Space (1999). The European initiatives and space mitigation standard are reviewed by Alby *et al.* 2004 [3].

Of course the International Space Station (ISS) will be the most heavily shielded spacecraft ever flown. Critical components (e.g., habitable compartments and high press tanks) will normally be able to withstand the impact of debris as large as 1 cm in diameter. ISS will also have manoeuvering capability to avoid hazardous objects.

The necessity of debris mitigation is illustrated in the ESA Space Debris Mitigation Handbook 2002 (Klinkrad *et al.*, 2004 [167]). Algorithms for the derivation of appropriate collision avoidance strategies are presented by Sánchez-Ortiz *et al.* (2004 [269]).

Bibliography

[1] Y. Abe and T. Matsui. Early evolution of the earth: Accretion, atmosphere formation, and thermal history. *Journal of Geophys. Research*, 91:291, September 1986.

[2] Y. Abe and T. Matsui. Evolution of an Impact-Generated H_2O-CO_2 Atmosphere and Formation of a Hot Proto-Ocean on Earth. *Journal of Atmospheric Sciences*, 45:3081–3101, November 1988.

[3] F. Alby, D. Alwes, L. Anselmo, H. Baccini, C. Bonnal, R. Crowther, W. Flury, R. Jehn, H. Klinkrad, C. Portelli, and R. Tremayne-Smith. The European Space Debris Safety and Mitigation Standard. *Advances in Space Research*, 34:1260–1263, 2004.

[4] V. A. Alexeev and G. K. Ustinova. Meteorite Data on the Solar Modulation of Galactic Cosmic Rays and an Inference on the Solar Activity Influence on Climate of the Earth. In S. Mackwell and E. Stansbery, editors, *36th Annual Lunar and Planetary Science Conference*, page 1012, March 2005.

[5] R. B. Alley and P. U. Clark. The Deglaciation of the Northern Hemisphere: A Global Perspective. *Annual Review of Earth and Planetary Sciences*, 27:149–182, 1999.

[6] L. W. Alvarez, W. Alvarez, F. Asaro, and H. V. Michel. Extraterrestrial Cause for the Cretaceous Tertiary Extinction. *Science*, 208:1095, 1980.

[7] E. I. Alves and A. R. Baptista. Rock Magnetic Fields Shield the Surface of Mars from Harmful Radiation. In S. Mackwell and E. Stansbery, editors, *Lunar and Planetary Institute Conference Abstracts*, page 1540, March 2004.

[8] D. M. Amron and R. L. Moy. Stratospheric ozone depletion and its relationship to skin cancer. *Journal of Dermatoligic Surgery and Oncology*, 17:370–372, 1991.

[9] L. Andersson, L. Eliassson, and P. Wintoft. In *Proccedings of the Workshop on Space Weather, 11-13 November 1998, ESTEC, Noordwijk WP-155*, 1999.

[10] A. L. Andrady, S. H. Hamid, X. Hu, and A. Torikai. Effects of increased solar ultraviolet radiation on materials. *Journal of Photochemistry and Photobiology B*, 1998.

[11] J. K. Angell and J. Korshover. *Month Weather Rev.*, 1976.

[12] J. Arnaud, C. Briand, and G. Ceppatelli. First observational campaign at the THEMIS: image quality and seeing. *New Astronomy Review*, 42:499–501, November 1998.

[13] C. Arpesella. In *BOREXINO Prop.*, 1992.

[14] M. J. Aschwanden, A. I. Poland, and D. M. Rabin. The New Solar Corona. *Annual Review of Astronomy and Astrophysics*, 39:175–210, 2001.

[15] S. Asiati, W. Sinambela, and R. Hidayati. Stratospheric, Tropospheric and Total Ozone Variation Correlated with Solar Activity. In *35th COSPAR Scientific Assembly*, page 966, 2004.

[16] W. Atwell, P. Saganti, F. A. Cucinotta, and C. J. Zeitlin. A space radiation shielding model of the Martian radiation environment experiment (MARIE). *Advances in Space Research*, 33:2219–2221, 2004.

[17] Behr A. Bruzek A. Durrant C. J. Enslin H. Fechtig H. Fricke W. Gondolatsch F. Gruen H. Hachenberg O. Ip W. H. Jessberger E. K. Kristen T. Leinert Ch. Lembke D. Palme H. Philipp W. Rahe J. Schmahl G. Scholer M. Schubart J. Solf J. Staubert R. Suess H. E. Truemper J. Weigelt G. West R. M. Wolf R. Zeh D. Axford, W. I. *Landolt-Börnstein, Numerical Data and Fundamental Relationships in Science and Technology, Vol 2a*. Springer, 1981.

[18] W. I. Axford. The origins of high-energy cosmic rays. *Astrophysical Journal Supplement*, 90:937–944, February 1994.

[19] R. B. Bacastow, C. D. Keeling, and T. P. Whorf. Seasonal amplitude increase in atmospheric CO2 concentration at Mauna Loa, Hawaii, 1959-1982. *Journal of Geophys. Research*, 90:10529, October 1985.

[20] D. Badhwar, G. *Radiat. Res.*, 148, 1997.

[21] J. N. Bahcall and R. J. Davis. The Evolution of Neutrino Astronomy. *Publ. Astr. Soc. of the Pacific*, 112:429–433, April 2000.

[22] J. N. Bahcall and M. H. Pinsonneault. Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Reviews of Modern Physics*, 64:885–926, 1992.

[23] J. N. Bahcall, M. H. Pinsonneault, S. Basu, and J. Christensen-Dalsgaard. Are Standard Solar Models Reliable? *Physical Review Letters*, 78:171–174, January 1997.

[24] T. Bai. High Flare Activity in the Late Declining Phase of Cycle 23. *Solar Physics*, 234:409–419, April 2006.

[25] S. Baliunas and W. Soon. Are Variations in the Length of the Activity Cycle Related to Changes in Brightness in Solar-Type Stars? *Astrophysical Journal*, 450:896, September 1995.

[26] J. R. Barnes and A. Collier Cameron. Starspot patterns on the M dwarfs HK Aqr and RE 1816 +541. *Monthly Notices*, 326:950–958, September 2001.

[27] R. E. Basher. Basic science of solar radiation and its ultraviolet components. In *Proceedings of the seminar on solar ultraviolet radiation, R. Basher (Ed), NZMS,*, 1981.

[28] D. Basu. Solar neutrino in relation to solar activity. *Solar Physics*, 142:205–208, November 1992.

[29] D. Basu. Radius of the Sun in relation to solar activity. *Solar Physics*, 183:291–294, December 1998.

[30] J. M. Beckers and T. M. Brown. *Oss. Mem. d. Oss. Astrofis. d. Arcetri*, 106:189, 1978.

[31] T. R. Bedding and H. Kjeldsen. Observing solar-like oscillations: recent results. *Memorie della Societa Astronomica Italiana*, 77:384, 2006.

[32] J. Beer. Long-term indirect indices of solar variability. *Space Science Reviews*, 94:53–66, November 2000.

[33] A. O. Benz. Coronal Heating and Microflares in Solar Active and Quiet Regions. *Astronomische Nachrichten Supplement*, 324:8, July 2003.

[34] A. Berger. The Milankovitch astronomical theory of paleoclimates: A modern review. *Vistas in Astronomy*, 24:103–122, 1980.

[35] R. A. Berner, A. C. Lasaga, and R. M Garrels. *Amer. J. Sci.*, 283:641–683, 1983.

[36] P. A. M. Berry. Periodicities in the sunspot cycle. *Vistas in Astronomy*, 30:97–108, 1987.

[37] J. Birn, T. G. Forbes, E. W. Hones, Jr., S. J. Bame, and G. Paschmann. On the velocity distribution of ion jets during substorm recovery. *Journal of Geophys. Research*, 86:9001–9006, October 1981.

[38] P. R. Boberg and A. J. Tylka. Variation in Solar Energetic Particle Elemental Composition Observedby ACE and Wind. In *AIP Conf. Proc. 528: Acceleration and Transport of Energetic Particles Observed in the Heliosphere*, page 115, 2000.

[39] P. Bochsler. Solar Wind Composition at Solar Maximum. *Space Science Reviews*, 97:113–121, 2001.

[40] T. J. Bogdan. Sunspot Oscillations: A Review - (Invited Review). *Solar Physics*, 192:373–394, March 2000.

[41] T. J. Bogdan, M. Carlsson, V. H. Hansteen, A. McMurry, C. S. Rosenthal, M. Johnson, S. Petty-Powell, E. J. Zita, R. F. Stein, S. W. McIntosh, and Å. Nordlund. Waves in the Magnetized Solar Atmosphere. II. Waves from Localized Sources in Magnetic Flux Concentrations. *Astrophysical Journal*, 599:626–660, December 2003.

[42] B. F. Bohor. In *Global catastrophes in Earth history, V.L. Sharpton*, 1990.

[43] J. A. Bonet. High Spatial Resolution Imaging in Solar Physics. In *ASSL Vol. 239: Motions in the Solar Atmosphere*, pages 1–12, 1999.

[44] M. Boscherini, O. Adriani, M. Bongi, L. Bonechi, G. Castellini, R. D'Alessandro, A. Gabbanini, M. Grandi, W. Menn, P. Papini, S. B. Ricciarini, M. Simon, P. Spillantini, S. Straulino, F. Taccetti, M. Tesi, and E. Vannuccini. Radiation damage of electronic components in space environment. *Nuclear Instruments and Methods in Physics Research A*, 514:112–116, November 2003.

[45] D. H. Boteler, R. J. Pirjola, and H. Nevanlinna. The effects of geomagnetic disturbances on electrical systems at the earth's surface. *Advances in Space Research*, 22:17–27, 1998.

[46] J. Bourgeois, P. L. Wiberg, and T. A. Hansen. Sedimentological Effects of Tsunamis, with Particular Reference to Impact-Generated and Volcanogenic Waves. *LPI Contributions*, 673:21, 1988.

[47] P. A. Bradley. The ionosphere and radio communications. In *ICC '84 - Links for the future: Science, systems and services for communications; Proceedings of the International Conference on Communications, Amsterdam, Netherlands, May 14-17, 1984. Volume 1 (A85-36626 16-32). New York/Amsterdam, Institute of Electrical and Electronics Engineers, Inc./North-Holland, 1984, p. 171-174.*, pages 171–174, 1984.

[48] P. D. Bradley and E. Normand. *IEEE Trans. Nucl. Sci.*, 45:6, 1998.

[49] L. Burlaga, E. Sittler, F. Mariani, and R. Schwenn. Magnetic loop behind an interplanetary shock - Voyager, Helios, and IMP 8 observations. *Journal of Geophys. Research*, 86:6673–6684, August 1981.

[50] L. F. Burlaga. Magnetic fields and plasmas in the inner heliosphere: Helios results. *Planetary and Space Science*, 49:1619–1627, December 2001.

[51] V. A. Burov. On the possibility of getting economically sound forecasts of rare space weather events. *Advances in Space Research*, 37:1247–1250, 2006.

[52] L. B. Callis, J. C. Alpert, and M. A. Geller. An assessment of thermal, wind, and planetary wave changes in the middle and lower atmosphere due to 11-year UV flux variations. *Journal of Geophys. Research*, 90:2273–2282, February 1985.

[53] R. A. Calvo, H. A. Ceccato, and R. D. Piacentini. Neural network prediction of solar activity. *Astrophysical Journal*, 444:916–921, May 1995.

[54] H. V. Cane and W. C. Erickson. Studies of Space Weather Using Solar Radio Bursts. In N. Kassim, M. Perez, W. Junor, and P. Henning, editors, *Astronomical Society of the Pacific Conference Series*, page 133, January 2006.

[55] D. C. Catling, K. J. Zahnle, and C. P. McKay. Biogenic Methane, Hydrogen Escape, and the Irreversible Oxidation of Early Earth. *Science*, 293:839–843, August 2001.

[56] F. Cattaneo. The Solar Dynamo Problem. In *ASSL Vol. 225: SCORe'96: Solar Convection and Oscillations and their Relationship*, pages 201–222, December 1997.

[57] F. Cattaneo. Dynamo Theory and the Origin of Small Scale Magnetic Fields. In *ASSL Vol. 239: Motions in the Solar Atmosphere*, pages 119–137, 1999.

[58] F. Cattaneo, D. Lenz, and N. Weiss. On the Origin of the Solar Mesogranulation. *Astrophysical Journal Letters*, 563:L91–L94, December 2001.

[59] G. Cauzzi, A. Falchi, and R. Falciani. Network and internetwork: a compared multiwavelength analysis. *Astronomy and Astrophysics*, 357:1093–1104, May 2000.

[60] S. Cecchini, G. Giacomelli, D. Haşegan, G. Mandrioli, O. Mariş, L. Patrizii, A. Plaian, V. Popa, L. Ştefanov, and V. Văleanu. Search for neutrino decay during the 1999 solar eclipse. *Astrophysics and Space Science*, 273:35–41, September 2000.

[61] D. K. Chakrabarty and P. Chakrabarty. The evolution of ozone with changing solar activity. *Geophysical Research Letter*, 9:76–78, January 1982.

[62] S. Chandrasekhar. A Theorem on Rotating Polytropes. *Astrophysical Journal*, 134:662–664, September 1961.

[63] C. R. Chapman and D. Morrison. Impacts on the Earth by asteroids and comets: assessing the hazard. *Nature*, 367:33–40, January 1994.

[64] Morozov N. Johnson N. McKnight D. Maclay Chernyavskiy, G. *IAA-94-IAA6.5.696*, 1994.

[65] A. R. Choudhuri, H. Auffret, and E. R. Priest. Implications of rapid footpoint motions of photospheric flux tubes for coronal heating. *Solar Physics*, 143:49–68, January 1993.

[66] J. Christensen-Dalsgaard. An Introduction to Solar Oscillations and Helioseismology. In V. Celebonovic, D. Gough, and W. Däppen, editors, *AIP Conf. Proc. 731: Equation-of-State and Phase-Transition in Models of Ordinary Astrophysical Matter*, pages 18–46, November 2004.

[67] J. Christensen-Dalsgaard, W. Dappen, S. V. Ajukov, E. R. Anderson, H. M. Antia, S. Basu, V. A. Baturin, G. Berthomieu, B. Chaboyer, S. M. Chitre, A. N. Cox, P. Demarque, J. Donatowicz, W. A. Dziembowski, M. Gabriel, D. O. Gough, D. B. Guenther, J. A. Guzik, J. W. Harvey, F. Hill, G. Houdek, C. A. Iglesias, A. G. Kosovichev, J. W. Leibacher, P. Morel, C. R. Proffitt, J. Provost, J. Reiter, E. J. Rhodes, F. J. Rogers, I. W. Roxburgh, M. J. Thompson, and R. K. Ulrich. The Current State of Solar Modeling. *Science*, 272:1286, 1996.

[68] R. T. Clancy and D. W. Rusch. Climatology and trends of mesospheric (58-90) temperatures based upon 1982-1986 SME limb scattering profiles. *Journal of Geophys. Research*, 94:3377–3393, March 1989.

[69] C. S. Cockell. Ultraviolet Radiation and the Photobiology of Earth's Early Oceans. *Origins of Life and Evolution of the Biosphere*, 30:467–500, October 2000.

[70] T. J. Crowley. Causes of Climate Change Over the Past 1000 Years. *Science*, 289:270–277, July 2000.

[71] R. D. D'Arrigo and B. C. Jacoby. *Climatic Change*, 1993.

[72] D. S. P. Dearborn and J. B. Blake. Magnetic Fields and the Solar Constant. *Nature*, 287:365, September 1980.

[73] D. S. P. Dearborn and M. J. Newman. Efficiency of convection and time variation of the solar constant. *Science*, 201:150, July 1978.

[74] M. Di Martino, M. Delbò, L. Saba, A. Cellino, V. Zappalà, S. Montebugnoli, S. Righini, L. Zoni, R. Orosei, and F. Tosi. The SRT as radar for asteroid and space debris studies. *Memorie della Societa Astronomica Italiana Supplement*, 10:180, 2006.

[75] B. Diffey. TOPICAL REVIEW: Climate change, ozone depletion and the impact on ultraviolet exposure of human skin. *Physics in Medicine and Biology*, 49:1, January 2004.

[76] M. Dikpati, G. de Toma, and P. A. Gilman. Predicting the strength of solar cycle 24 using a flux-transport dynamo-based tool. *Geophysical Research Letter*, 33:5102, March 2006.

[77] E. Doornbos and H. Klinkrad. Modelling of space weather effects on satellite drag. *Advances in Space Research*, 37:1229–1239, 2006.

[78] L. I. Dorman. Estimation of long-term cosmic ray intensity variation in near future and prediction of their contribution in expected global climate change. *Advances in Space Research*, 35:496–503, 2005.

[79] D. Dravins, L. Lindegren, A. Nordlund, and D. A. Vandenberg. The distant future of solar activity: A case study of Beta Hydri. I - Stellar evolution, lithium abundance, and photospheric structure. *Astrophysical Journal*, 403:385–395, January 1993.

[80] D. Dravins, L. Lindegren, A. Nordlund, and D. A. Vandenberg. The distant future of solar activity: A case study of Beta Hydri. I - Stellar evolution, lithium abundance, and photospheric structure. *Astrophysical Journal*, 403:385–395, January 1993.

[81] J. R. Dudeney and R. I. Kressman. Empirical models of the electron concentration of the ionosphere and their value for radio communications purposes. *Radio Science*, 21:319–330, June 1986.

[82] J. W. Dungey. Interplanetary Magnetic Field and the Auroral Zones. *Physical Review Letters*, 6:47–48, January 1961.

[83] T. L. Duvall. a New Method to Search for Solar Gravity Mode Oscillations. In D. Danesy, editor, *ESA SP-559: SOHO 14 Helio- and Asteroseismology: Towards a Golden Future*, page 412, October 2004.

[84] T. L. Duvall, A. G. Kosovichev, P. H. Scherrer, R. S. Bogart, R. I. Bush, C. de Forest, J. T. Hoeksema, J. Schou, J. L. R. Saba, T. D. Tarbell, A. M. Title, C. J. Wolfson, and P. N. Milford. Time-Distance Helioseismology with the MDI Instrument: Initial Results. *Solar Physics*, 170:63–73, 1997.

[85] C. Dyer. In *SOLSPA 2001, ESA SP Ser 477*, 2001.

[86] W. A. Dziembowski, P. R. Goode, A. G. Kosovichev, and J. Schou. Signatures of the Rise of Cycle 23. *Astrophysical Journal*, 537:1026–1038, July 2000.

[87] J. A. Eddy. The Maunder Minimum. *Science*, 192:1189–1202, June 1976.

[88] J. A. Eddy. The Langley Years. *Bulletin of the American Astronomical Society*, 20:949, September 1988.

[89] M. Emilio, J. R. Kuhn, R. I. Bush, and P. Scherrer. On the Constancy of the Solar Diameter. *Astrophysical Journal*, 543:1007–1010, November 2000.

[90] O. Espagnet, R. Muller, T. Roudier, N. Mein, and P. Mein. Penetration of the solar granulation into the photosphere: height dependence of intensity and velocity fluctuations. *Astronomy and Astrophysics Supplement*, 109:79–108, January 1995.

[91] O. Espagnet, R. Muller, T. Roudier, P. Mein, N. Mein, and J. M. Malherbe. Spatial relation between the 5-minute oscillations and granulation patterns. *Astronomy and Astrophysics*, 313:297–305, September 1996.

[92] D. M. Etheridge, G. I. Pearman, and F. de Silva. Atmospheric trace-gas variations as revealed by air trapped in an ice core from Law Dome, Antarctica. *Annals of Glaciology, vol.10, pp.28-33*, 10:28–33, 1988.

[93] D. M. Etheridge, L. P. Steele, R. J. Langenfelds, R. L.and Francey, J.-M. Barnola, and V. I. Morgan. Natural and anthropogenic changes in atmospheric CO_2 over the last 1000 years from air in Antarctic ice and firn. *Journal of Geophys. Research*, 101:4115–4128, 1996.

[94] Starr R. Brückner J. Boynton W. V. Bailey S. H. Tromka J. L. Evans, L. G. *Nuclear Instruments and Methods in Physics Research, Sec. A.*, 422:586, 1999.

[95] D. H. Fairfield and L. J. Cahill, Jr. Transition Region Magnetic Field and Polar Magnetic Disturbances. *Journal of Geophys. Research*, 71:155–+, January 1966.

[96] McLain D. R. Favorite, F. *Nature*, 244:139, 1973.

[97] G. H. Fisher, Y. Fan, D. W. Longcope, M. G. Linton, and A. A. Pevtsov. The Solar Dynamo and Emerging Flux - (Invited Review). *Solar Physics*, 192:119–139, March 2000.

[98] B. Fleck. Highlights from SOHO and Future Space Missions. In *The Dynamic Sun, Proceedings of the Summerschool and Workshop held at the Solar Observatory*, page 1, May 2001.

[99] A. Fossum and M. Carlsson. High-frequency acoustic waves are not sufficient to heat the solar chromosphere. *Nature*, 435:919–921, June 2005.

[100] P. Foukal and J. Lean. Magnetic modulation of solar luminosity by photospheric activity. *Astrophysical Journal*, 328:347–357, May 1988.

[101] P. Foukal and J. Lean. An empirical model of total solar irradiance variation between 1874 and 1988. *Science*, 247:556–558, February 1990.

[102] C. Fröhlich. Observations of Irradiance Variations. *Space Science Reviews*, 94:15–24, November 2000.

[103] E. Friis-Christensen and K. Lassen. Length of the solar cycle: an indicator of solar activity closely associated with climate. *Science*, 254:698–700, November 1991.

[104] C. Froehlich. Variability of the solar 'constant' on time scales of minutes to years. *Journal of Geophys. Research*, 92:796–800, January 1987.

[105] C. Frohlich, J. Romero, H. Roth, C. Wehrli, B. N. Andersen, T. Appourchaux, V. Domingo, U. Telljohann, G. Berthomieu, P. Delache, J. Provost, T. Toutain, D. A. Crommelynck, A. Chevalier, A. Fichot, W. Dappen, D. Gough, T. Hoeksema, A. Jimenez, M. F. Gomez, J. M. Herreros, T. R. Cortes, A. R. Jones, J. M. Pap, and R. C. Willson. VIRGO: Experiment for Helioseismology and Solar Irradiance Monitoring. *Solar Physics*, 162:101–128, 1995.

[106] V. Gaizauskas. Preflare activity. *Solar Physics*, 121:135–152, 1989.

[107] J.-C. Gerard. Modelling the Climatic Response to Solar Variability. *Royal Society of London Philosophical Transactions Series A*, 330:561–573, April 1990.

[108] R. Gertsch, L. S. Gertsch, and J. L. Remo. Mining Near-Earth Resources. In *Near-Earth Objects*, page 511, January 1997.

[109] S. E. Gibson. Global Solar Wind Structure from Solar Minimum to Solar Maximum: Sources and Evolution. *Space Science Reviews*, 97:69–79, 2001.

[110] R. L. Gilliland. Solar Luminosity Variations - Models of Solar Convection Zone. *Nature*, 286:838, August 1980.

[111] R. L. Gilliland. Solar radius variations over the past 265 years. *Astrophysical Journal*, 248:1144–1155, September 1981.

[112] P. A. Gilman. The tachocline and the solar dynamo. *Astronomische Nachrichten*, 326:208–217, 2005.

[113] H. Gleisner and H. Lundstedt. A neural network-based local model for prediction of geomagnetic disturbances. *Journal of Geophys. Research*, pages 8425–8434, May 2001.

[114] H. Gleisner and J. Watermann. Concepts of medium-range (1 3 days) geomagnetic forecasting. *Advances in Space Research*, 37:1116–1123, 2006.

[115] P. Goldreich, N. Murray, and P. Kumar. Excitation of solar p-modes. *Astrophysical Journal*, 424:466–479, March 1994.

[116] P. R. Goode, L. H. Strous, T. R. Rimmele, and R. T. Stebbins. On the Origin of Solar Oscillations. *Astrophysical Journal Letters*, 495:L27+, March 1998.

[117] J. T. Gosling, E. Hildner, R. M. MacQueen, R. H. Munro, A. I. Poland, and C. L. Ross. The speeds of coronal mass ejection events. *Solar Physics*, 48:389–397, June 1976.

[118] D. O. Gough. Solar interior structure and luminosity variations. *Solar Physics*, 74:21–34, November 1981.

[119] D. O. Gough, A. G. Kosovichev, J. Toomre, E. Anderson, H. M. Antia, S. Basu, B. Chaboyer, S. M. Chitre, J. Christensen-Dalsgaard, W. A. Dziembowski, A. Eff-Darwich, J. R. Elliott, P. M. Giles, P. R. Goode, J. A. Guzik, J. W. Harvey, F. Hill, J. W. Leibacher, M. J. P. F. G. Monteiro, O. Richard, T. Sekii, H. Shibahashi, M. Takata, M. J. Thompson, S. Vauclair, and S. V. Vorontsov. The Seismic Structure of the Sun. *Science*, 272:1296, 1996.

[120] G. A. Graham, A. Sexton, M. M. Grady, and I. P. Wright. Further attempts to constrain the nature of the impact residues in the HST solar array panels. *Advances in Space Research*, 20:1461–1465, 1997.

[121] E. F. Guinan and I. Ribas. Our Changing Sun: The Role of Solar Nuclear Evolution and Magnetic Activity on Earth's Atmosphere and Climate. In *ASP Conf. Ser. 269: The Evolving Sun and its Influence on Planetary Environments*, page 85, 2002.

[122] J. D. Haigh. The Role of Stratospheric Ozone in Modulating the Solar Radiative Forcing of Climate. *Nature*, 370:544, August 1994.

[123] J. D. Haigh. *Science*, 272:981, 1996.

[124] B. Haisch, K. T. Strong, and M. Rodono. Flares on the sun and other stars. *Annual Review of Astronomy and Astrophysics*, 29:275–324, 1991.

[125] J. E. Hansen and A. A. Lacis. Sun and Dust Versus Greenhouse Gases - an Assessment of Their Relative Roles in Global Climate Change. *Nature*, 346:713, August 1990.

[126] C. Hanuise, J. C. Cerisier, F. Auchère, K. Bocchialini, S. Bruinsma, N. Cornilleau-Wehrlin, N. Jakowski, C. Lathuillère, M. Menvielle, J.-J. Valette, N. Vilmer, J. Watermann, and P. Yaya. From the Sun to the Earth: impact of the 27-28 May 2003 solar events on the magnetosphere, ionosphere and thermosphere. *Annales Geophysicae*, 24:129–151, January 2006.

[127] C. K. Hargrove and D. J. Paterson. Solar-neutrino neutral-current detection methods in the Sudbury Neutrino Observatory. *Canadian Journal of Physics*, 69:1309–1316, November 1991.

[128] D. H. Hathaway, J. G. Beck, R. S. Bogart, K. T. Bachmann, G. Khatri, J. M. Petitto, S. Han, and J. Raymond. The Photospheric Convection Spectrum. *Solar Physics*, 193:299–312, April 2000.

[129] W. C. Haxton. *Neutrino oscillations and the solar neutrino problem*, page 65. Current aspects of neutrino physics, 2001.

[130] G. R. Heckman. Solar proton event forecasts. In *Interplanetary Particle Environment*, pages 91–100, April 1988.

[131] P. Heinzel and U. Anzer. Physics of solar prominences. In *Solar Magnetic Phenomena, Proceedings of the 3rd Summerschool and Workshop held at the Solar Observatory Kanzelhöhe, Kärnten, Austria, August 25 – September 5, 2003. Edited by A. Hanslmeier, A. Veronig, and M. Messerotti. Astronomy and Astrophysics Space Science Library, vol. 320, Springer, Dordrecht, The Netherlands, 2005., p.115-138*, pages 115–138, 2005.

[132] J. V. Hernandez, T. Tajima, and W. Horton. Neural net forecasting for geomagnetic activity. *Geophysical Research Letter*, 20:2707–2710, December 1993.

[133] J. G. Hills and Ch. L. Mader. *Ann. New York Acad. of Sciences*, 822:381, 1997.

[134] C. O. Hines. A Possible Mechanism for the Production of Sun-Weather Correlations. *Journal of Atmospheric Sciences*, 31:589–591, March 1974.

[135] N. M. Hoekzema, R. J. Rutten, P. N. Brandt, and R. A. Shine. Small-scale topology of solar atmosphere dynamics. I. Wave sources and wave diffraction. *Astronomy and Astrophysics*, 329:276–290, January 1998.

[136] L. L. Hood. Coupled stratospheric ozone and temperature responses to short-term changes in solar ultraviolet flux - An analysis of Nimbus 7 SBUV and SAMS data. *Journal of Geophys. Research*, 91:5264–5276, April 1986.

[137] L. L. Hood. Solar ultraviolet radiation induced variations in the stratosphere and mesosphere. *Journal of Geophys. Research*, 92:876–888, January 1987.

[138] Jenkins G. J. Houghton, J. T. and J. J. Ephraums. *Climate Change: The Ipcc Scientific Assessment*. Cambridge University Press, 1990, 1990.

[139] T. A. Howard, D. F. Webb, S. J. Tappin, D. R. Mizuno, and J. C. Johnston. Tracking halo coronal mass ejections from 0-1 AU and space weather forecasting using the Solar Mass Ejection Imager (SMEI). *Journal of Geophysical Research (Space Physics)*, 111:4105, April 2006.

[140] D. V. Hoyt and K. H. Schatten. A discussion of plausible solar irradiance variations, 1700-1992. *Journal of Geophys. Research*, 98:18895, November 1993.

[141] D. V. Hoyt and K. H. Schatten. *The role of the sun in climate change*. The role of the sun in climate change / Douglas V. Hoyt, Kenneth H. Schatten. New York : Oxford University Press, 1997. QC883.2.S6 H69 1997, 1997.

[142] M. K. Hughes and H. F. Diaz, editors. *The medieval warm period*, 1994.

[143] D. M. Hunten, G. Cremonese, A. L. Sprague, R. E. Hill, S. Verani, and R. W. H. Kozlowski. The Leonid Meteor Shower and the Lunar Sodium Atmosphere. *Icarus*, 136:298–303, December 1998.

[144] N. E. Hurlburt. Solar Magnetoconvection. In *SOHO-9 Workshop on Helioseismic Diagnostics of Solar Convection and Activity*, 1999.

[145] J. Hyde and E. Christiansen. Space Shuttle Meteoroid and Orbital Debris Threat Assessment Procedure. In *COSPAR, Plenary Meeting*, 2002.

[146] T. T. Ishii, H. Kurokawa, and T. T. Takeuchi. Characteristics of Flare-productive Sunspot Groups. *The Sun and Space Weather, 24th meeting of the IAU, Joint Discussion 7, August 2000, Manchester, England.*, 7, 2000.

[147] B. V. Jackson, P. P. Hick, and A. Buffington. Halo CME's - Will They Hit or Miss Earth? *AGU Fall Meeting Abstracts*, pages A474+, December 2002.

[148] J. M. Jensen, H. Lundstedt, M. J. Thompson, F. P. Pijpers, and S. P. Rajaguru. Application of local-area helioseismic methods as predictors of spaceweather. In *35th COSPAR Scientific Assembly*, page 2195, 2004.

[149] S. J. Johnson, W. Dalsgaard, H. B. Clausen, and C. C. Langway. *Nature*, 227:482–483, 1970.

[150] A. M. Jorgensen and H. Karimabadi. Understanding space weather via analysis, prediction, modeling, and classification supported by Genetic Programming. *AGU Fall Meeting Abstracts*, pages A311+, December 2005.

[151] W. Kalkofen. The heating of the quiet solar chromosphere. In *IAU Symp. 142: Basic Plasma Processes on the Sun*, pages 197–204, 1990.

[152] W. Kalkofen. Oscillations in Chromospheric Network Bright Points. *Astrophysical Journal Letters*, 486:L145+, September 1997.

[153] J. F. Kasting. Habitable Zones around Low Mass Stars and the Search For Extraterrestrial Life. *Origins of Life and Evolution of the Biosphere*, 27:291–3107, 1997.

[154] J. F. Kasting. Evolution of Earth's Atmosphere and Climate. *AGU Fall Meeting Abstracts*, December 2004.

[155] J. F. Kasting and O. B. Toon. *Climate evolution on the terrestrial planets*, pages 423–449. Origin and Evolution of Planetary and Satellite Atmospheres, 1989.

[156] G. M. Keating. The response of ozone to solar activity variations - A review. *Solar Physics*, 74:321–347, 1981.

[157] G. M. Keating, M. C. Pitts, G. Brasseur, and A. de Rudder. Response of middle atmosphere to short-term solar ultraviolet variations. I - Observations. II - Theory. *Journal of Geophys. Research*, 92:889–914, January 1987.

[158] L. P. Keller, S. J. Wentworth, J. Gezo, D. S. McKay, L. A. Taylor, C. Pieters, and R. V. Morris. Space Weathering Alteration of Lunar Soil Grains. In *Lunar and Planetary Institute Conference Abstracts*, page 1820, March 1999.

[159] R. A. Kerr. Sunspot Weather Link Holding UP. *Science*, 242:1124, November 1988.

[160] E. V. Khomenko, R. I. Kostik, and N. G. Shchukina. Five-minute oscillations above granules and intergranular lanes. *Astronomy and Astrophysics*, 369:660–671, April 2001.

[161] J. Kiefer. *Phys. Med. Suppl*, 17:1, 2001.

[162] R.-S. Kim, K.-S. Cho, Y.-J. Moon, Y.-H. Kim, Y. Yi, M. Dryer, S.-C. Bong, and Y.-D. Park. Forecast evaluation of the coronal mass ejection (CME) geoeffectiveness using halo CMEs from 1997 to 2003. *Journal of Geophysical Research (Space Physics)*, 110:11104, November 2005.

[163] R. Kippenhahn and A. Schlüter. Eine Theorie der solaren Filamente. Mit 7 Textabbildungen. *Zeitschrift fur Astrophysics*, 43:36, 1957.

[164] D. Kiselman, R. J. Rutten, and B. Plez. The formation of G-band bright points I: Standard LTE modelling. In *IAU Symposium*, page 287, 2001.

[165] H. Kjeldsen, T. R. Bedding, R. P. Butler, J. Christensen-Dalsgaard, L. L. Kiss, C. McCarthy, G. W. Marcy, C. G. Tinney, and J. T. Wright. Solar-like Oscillations in α Centauri B. *Astrophysical Journal*, 635:1281–1290, December 2005.

[166] A. Klassen, S. Pohjolainen, and K.-L. Klein. Type II radio precursor and X-ray flare emission. *Solar Physics*, 218:197–210, December 2003.

[167] H. Klinkrad, P. Beltrami, S. Hauptmann, C. Martin, H. Sdunnus, H. Stokes, R. Walker, and J. Wilkinson. The ESA Space Debris Mitigation Handbook 2002. *Advances in Space Research*, 34:1251–1259, 2004.

[168] F. Kneer and M. von Uexküll. Diagnostics and Dynamics of the Solar Chromosphere. In *ASSL Vol. 239: Motions in the Solar Atmosphere*, pages 99–118, 1999.

[169] D. J. Knipp, W. K. Tobiska, and B. A. Emery. Direct and Indirect Thermospheric Heating Sources for Solar Cycles 21 23. *Solar Physics*, 224:495–505, October 2004.

[170] Y. Kobayashi, H. Watanabe, M. Kikuchi, and I. Narumi. Effect of the Space Environment on the Induction of DNA-repair Related Proteins and Recovery from Radiation Damage. *Advances in Space Research*, 25:2103–2106, 2000.

[171] A. G. Kosovichev, J. Schou, P. H. Scherrer, P. H. Goode, W. A. Dziembowski, E. J. Rhodes, and The SOI Structure Inversion Team. Spherical and aspherical structure of the sun: First year of SOHO/MDI observations. In *IAU Symp. 185: New Eyes to See Inside the Sun and Stars*, page 157, 1998.

[172] V. A. Krasnopolsky. Some problems related to the origin of methane on Mars. *Icarus*, 180:359–367, February 2006.

[173] P. I. Krastev and A. Y. Smirnov. Global analysis with SNO: Toward the solution of the solar neutrino problem. *Physical Review D*, 65(7):073022, April 2002.

[174] F. Krause and K. H. Raedler. *Mean-field magnetohydrodynamics and dynamo theory*. Oxford: Pergamon Press, 1980, 1980.

[175] A. S. Krieger, A. F. Timothy, and E. C. Roelof. A Coronal Hole and Its Identification as the Source of a High Velocity Solar Wind Stream. *Solar Physics*, 29:505, 1973.

[176] A. M. Krymskii, T. K. Breus, N. F. Ness, and M. H. AcuÑa. The IMF pile-up regions near the Earth and Venus: lessons for the solar wind - Mars interaction. *Space Science Reviews*, 92:535–564, May 2000.

[177] Y. Kubo. Position and radius of the Sun determined by solar eclipses in combination with lunar occultations. *Publications of the Astronomical Society of Japan*, 45:819–829, December 1993.

[178] K. Kudela and M. Storini. Possible tools for space weather issues from cosmic ray continuous records. *Advances in Space Research*, 37:1443–1449, 2006.

[179] K. Labitzke. Sunspots, the QBO, and the stratospheric temperature in the north polar region. *Geophysical Research Letter*, 14:535–537, 1987.

[180] K. Labitzke and H. van Loon. Some Aspects of a Probable Association Between Atmospheric Variability and the 11-year Solar Cycle. In *Solar Radiative Output Variation*, page 57, 1988.

[181] J. Lachance, C. Coïa, A. C. Fozza, G. Czeremuszkin, A. Houdayer, and M. R. Wertheimer. Radiation-induced degradation of polymeric spacecraft materials under protective oxide coatings. *Nuclear Instruments and Methods in Physics Research B*, 185:328–335, December 2001.

[182] F. Laclare, C. Delmas, J. P. Coin, and A. Irbah. Measurements and Variations of the Solar Diameter. *Solar Physics*, 166:211–229, July 1996.

[183] M. Lagache. New data on the kinetics of the dissolution of alkali feldspars at 200 C in CO 2 charged water. *Geochimica et Cosmochimica Acta*, 40:157–161, February 1976.

[184] G. S. Lakhina, S. Alex, B. T. Tsurutani, and W. D. Gonzalez. Research on Historical Records of Geomagnetic Storms. In K. Dere, J. Wang, and Y. Yan, editors, *IAU Symposium*, pages 3–15, 2005.

[185] H. H. Lamb. *Climate Present, Past and Future, Vol. 2 Climatic history and future*. Methuen, London, 1977.

[186] H. Lammer, W. Stumptner, G. J. Molina-Cuberos, S. J. Bauer, and T. Owen. Nitrogen isotope fractionation and its consequence for Titan's atmospheric evolution. *Planetary and Space Science*, 48:529–543, May 2000.

[187] K. Langhans, G. B. Scharmer, D. Kiselman, M. G. Löfdahl, and T. E. Berger. Inclination of magnetic fields and flows in sunspot penumbrae. *Astronomy and Astrophysics*, 436:1087–1101, June 2005.

[188] P. Lantos. Predictions of Galactic Cosmic Ray Intensity Deduced from that of Sunspot Number. *Solar Physics*, 229:373–386, July 2005.

[189] J. Lean. The Sun's Variable Radiation and Its Relevance For Earth. *Annual Review of Astronomy and Astrophysics*, 35:33–67, 1997.

[190] J. L. Lean. Societal Impacts of Solar Electromagnetic Radiation. *Bulletin of the American Astronomical Society*, 32:840, May 2000.

[191] J. L. Lean, J. T. Mariska, K. T. Strong, H. S. Hudson, L. W. Acton, G. J. Rottman, T. N. Woods, and R. C. Willson. Correlated brightness variations in solar radiative output from the photosphere to the corona. *Geophysical Research Letter*, 22:655–658, March 1995.

[192] R. B. Leighton, R. W. Noyes, and G. W. Simon. Velocity Fields in the Solar Atmosphere. I. Preliminary Report. *Astrophysical Journal*, 135:474, March 1962.

[193] N. V. Leister and P. Benevides-Soares. Solar diameter variations. *Academie des Sciences Paris Comptes Rendus Serie Sciences Mathematiques*, 311:399–404, August 1990.

[194] R. P. Lepping and D. B. Berdichevsky. Magnetic Clouds as Intermediary Agents between their Solar Sources and the Magnetosphere. *AGU Spring Meeting Abstracts*, page 61, May 2001.

[195] M. R. Lewis, M. Carr, G. C. Feldman, W. Esaias, and C. McClain. *Nature*, 347:543, 1990.

[196] J. Lilensten and J. Bornarel. *Space Weather, Environment and Societies*. Space Weather, Environment and Societies. By Jean Lilensten, Researcher (CNRS), Planetary Laboratory of the Grenoble University, Université Joseph Fourier), France; Jean Bornarel, Professor Grenoble University (Université Joseph Fourier), Physics Spectrometry Laboratory, France, Springer Dordrecht, December 2005.

[197] B. W. Lites, R. J. Rutten, and T. E. Berger. Dynamics of the Solar Chromosphere. II. Ca II H_2V and K_2V Grains versus Internetwork Fields. *Astrophysical Journal*, 517:1013–1033, June 1999.

[198] C. Liu, J. Qiu, D. E. Gary, S. Krucker, and H. Wang. Studies of Microflares in RHESSI Hard X-Ray, Big Bear Solar Observatory Hα, and Michelson Doppler Imager Magnetograms. *Astrophysical Journal*, 604:442–448, March 2004.

[199] M. Lockwood and S. Foster. Long-Term Variations in the Magnetic Fields of the Sun and Possible Implications for Terrestrial Climate. In *ESA SP-463: The Solar Cycle and Terrestrial Climate, Solar and Space weather*, page 85, 2000.

[200] M. Lockwood, R. Stamper, and M. N. Wild. A doubling of the sun's coronal magnetic field during the past 100 years. *Nature*, 399:437–439, May 1999.

[201] J. E. Lovelock. *The Ages of Gaia*. Norton, New York, 1988.

[202] O.v.d. Lühe. In *The Dynamic Sun, Proceedings of the Summerschool and Workshop held at the Solar Observatory*, pages 43–68, May 2001.

[203] H. Lundstedt. The sun, space weather and GIC effects in Sweden. *Advances in Space Research*, 37:1182–1191, 2006.

[204] T. J. Lydon and S. Sofia. A Method for Incorporating the Effects of Large-Scale Magnetic Fields in the Study of Stellar Structure and Variability. *Astrophysical Journal Supplement*, 101:357, December 1995.

[205] W. B. Manchester, T. I. Gombosi, I. Roussev, A. Ridley, D. L. De Zeeuw, I. V. Sokolov, K. G. Powell, and G. Tóth. Modeling a space weather event from the Sun to the Earth: CME generation and interplanetary propagation. *Journal of Geophysical Research (Space Physics)*, 109:2107, February 2004.

[206] M. E. Mann, R. S. Bradley, and M. K. Hughes. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophysical Research Letter*, 26:759–762, 1999.

[207] T. Markvart. Radiation Damage in Solar Cells: a review . *Materials in Electronics*, 1990.

[208] T. Markvart, T. J. Cumberbatch, A. A. Dollery, and M. Walkden. Photon and electron degradation of boron-doped FZ silicon solar cells . In *Proc. Third European Symposium on Photovoltaic Generators in Space*, 1982.

[209] T. Markvart, A. F. W. Willoughby, and A. A. Goodbody, G. C.and Dollery. Defect interactions in silicon solar cells. In *Second Workshop on Radiation-Induced and/or Process-Related Electrically Active Defects in Semiconductors-Insulator Systems, North Carolina*, 1989.

[210] T. Markvart, A. F. W. Willoughby, G. C. Goodbody, and A. A. Dollery. Radiation-resistant silicon solar cell . In *Proc. 19th IEEE Photovoltaic Specialists Conference*, 1987.

[211] N. Marsh and H. Svensmark. Cosmic Rays, Clouds, and Climate. *Space Science Reviews*, 94:215–230, November 2000.

[212] N. D. Marsh and H. Svensmark. Low Cloud Properties Influenced by Cosmic Rays. *Physical Review Letters*, 85:5004–5007, December 2000.

[213] H. Mavromichalaki, G. Souvatzoglou, C. Sarlanis, G. Mariatos, C. Plainaki, M. Gerontidou, A. Belov, E. Eroshenko, and V. Yanke. Space weather prediction by cosmic rays. *Advances in Space Research*, 37:1141–1147, 2006.

[214] N. McBride and J. a. m. McDonnell. Meteoroid impacts on spacecraft:sporadics, streams, and the 1999 Leonids. *Planetary and Space Science*, 47:1005–1013, August 1999.

[215] B. M. McCormac and T. A. Seliga, editors. *Solar-terrestrial influences on weather and climate*, 1979.

[216] J. A. M. McDonnell and D. J. Gardner. Meteoroid Morphology and Densities: Decoding Satellite Impact Data. *Icarus*, 133:25–35, May 1998.

[217] J. D. McDougall. *Science*, 239:485, 1988.

[218] P. S. McIntosh. The classification of sunspot groups. *Solar Physics*, 125:251–267, February 1990.

[219] G. F. McKay, J. Dubeau, and I. Thomson. In *Proceedings of Spacebound 93 Conference, Canadian Space Agency, Ottawa*, 1993.

[220] R. L. McKenzie and J. M. Elwood. Intensity of solar ultraviolet radiation and its implications for skin cancer . *NZ Medical Journal*, 103:152, 1990.

[221] W. J. McNeil, S. T. Lai, and E. Murad. Charge Production due to Leonid Meteor Shower Impact on Spacecraft Surfaces. In *6th Spacecraft Charging Technology*, pages 187–191, November 1998.

[222] P. Mein. New Ground-Based Solar Instrumentation. *LNP Vol. 489: European Meeting on Solar Physics*, page 241, 1997.

[223] H. J. Melosh, N. M. Schneider, K. J. Zahnle, and D. Latham. Ignition of global wildfires at the Cretaceous/Tertiary boundary. *Nature*, 343:251–254, January 1990.

[224] I. Molotov, A. Konovalenko, G. Tuccari, V. Agapov, A. Antipenko, Y. Gorshenkov, A. Volvach, X. Liu, L. Litvinenko, I. Falkovich, O. Fedorov, S. Zasukha, V. Abrosimov, A. Pushkarev, M. Nechaeva, A. Dementiev, N. Dugin, V. Titenko, I. Shmeld, V. Jazykov, S. Buttaccio, C. Nicotra, A. Tsyukh, V. Nesteruk, and I. Puchinin. International Radar Space Debris Research. In D. Danesy, editor, *ESA SP-589: 4th European Conference on Space Debris*, page 83, August 2005.

[225] P. Morel, J. Provost, and G. Berthomieu. Updated solar models. *Astronomy and Astrophysics*, 327:349–360, November 1997.

[226] F. Moreno-Insertis. The magnetic field in the convection zone as a link between the active regions on the surface and the field in the solar interior. In *Solar Magnetic Fields*, pages 117–135, 1994.

[227] R. Muller. The Solar Granulation. In *ASSL Vol. 239: Motions in the Solar Atmosphere*, pages 35–70, 1999.

[228] R. Muller, J. C. Hulot, and T. Roudier. Perturbation of the granular pattern by the presence of magnetic flux tubes. *Solar Physics*, 119:229–243, 1989.

[229] U. Narain and P. Ulmschneider. Chromospheric and coronal heating mechanisms. *Space Science Reviews*, 54:377–445, December 1990.

[230] A. Neftel, E. Moor, H. Oeschger, and B. Stauffer. Evidence from polar ice cores for the increase in atmospheric CO2 in the past two centuries. *Nature*, 315:45–47, May 1985.

[231] W. M. Neupert. Comparison of Solar X-Ray Line Emission with Microwave Emission during Flares. *Astrophysical Journal Letters*, 153:L59+, July 1968.

[232] F. Noeel. Variations of the apparent solar semidiameter observed with the astrolabe of Santiago. *Astronomy and Astrophysics*, 325:825–827, September 1997.

[233] F. Noel. Observations of the Sun during 1994 with the astrolabe of Santiago. *Astronomy and Astrophysics Supplement*, 113:131, October 1995.

[234] A. Nordlund. Numerical simulations of the solar granulation. I - Basic equations and methods. *Astronomy and Astrophysics*, 107:1–10, March 1982.

[235] E. Normand. *IEEE Trans. Nucl. Sci.*, 1996.

[236] L. J. November, J. Toomre, K. B. Gebbie, and G. W. Simon. The detection of mesogranulation on the sun. *Astrophysical Journal Letters*, 245:L123–L126, May 1981.

[237] O. A. Olmedo, Jr., J. Zhang, H. Wechsler, K. Borne, and A. Poland. The Development of an Automatic Solar Eruptive Event Detection System (SEEDS). *AAS/Solar Physics Division Meeting*, 37:08.17, June 2006.

[238] S. Omori, K. Kitajima, and S. Maruyama. Subduction of the Carbonated Archean Oceanic-Crust and its Implication to Evolution of Earth's Early Atmosphere. *AGU Fall Meeting Abstracts*, pages B1043+, December 2003.

[239] M. Ossendrijver. The solar dynamo. *The Astronomy and Astrophysics Review*, 11:287–367, 2003.

[240] M. Ozima. Moon as a Recorder of Early Earth Evolution; A new Perspective. *AGU Fall Meeting Abstracts*, December 2005.

[241] M. Ozima, K. Seki, N. Terada, Y. N. Miura, F. A. Podosek, and H. Shinagawa. Terrestrial Atmospheric Components in Lunar Soils: Record of Early Earth Evolution. In *36th Annual Lunar and Planetary Science Conference*, page 1118, March 2005.

[242] B. E. Palle and C. J. Butler. . In *ESA SP-463*, 2000.

[243] J. Pap, M. Anklin, C. Fröhlich, C. Wehrli, F. Varadi, and L. Floyd. Variations in total solar and spectral irradiance as measured by the VIRGO experiment on SOHO. *Advances in Space Research*, 24:215–224, 1999.

[244] C. A Perry. *Int. J. of Climatology*, 14:969, 1994.

[245] J. W. Peters, T. Markvart, and A. F. W. Willoughby. A study of radiation-induced defects in silicon solar cells showing improved radiation resistance . *Materials Science Forum*, 1992.

[246] R. Pirjola, A. Viljanen, A. Pulkkinen, and O. Amm. Space weather risk in power systems and pipelines. *Physics and Chemistry of the Earth C*, 25:333–337, March 2000.

[247] M. S. Potgieter. The Modulation of Galactic Cosmic Rays in the Heliosphere: Theory and Models. *Space Science Reviews*, 83:147–158, January 1998.

[248] J. Powell, G. Maise, H. Ludewig, and M. Todosow. High-Performance Ultralight Nuclear Rockets for Near-Earth Objects Interaction Missions. In *Near-Earth Objects*, page 447, January 1997.

[249] E. R. Priest. Solar Flare Theory and the Status of Flare Understanding. In *ASP Conf. Ser. 206: High Energy Solar Physics Workshop - Anticipating Hess!*, page 13, 2000.

[250] A. Pulkkinen, S. Lindahl, A. Viljanen, and R. Pirjola. Geomagnetic storm of 29-31 October 2003: Geomagnetically induced currents and their relation to problems in the Swedish high-voltage power transmission system. *Space Weather*, 3:8, August 2005.

[251] M. P. Rast. The Scales of Granulation, Mesogranulation, and Supergranulation. *Astrophysical Journal*, 597:1200–1210, November 2003.

[252] D. M. Raup and J. J. Sepkoski. Testing for Periodicity of Extinction. *Science*, 241:94–96, July 1988.

[253] R. C. Reedy. Gamma Rays from Mercury Made by Solar Energetic Particles. *Bulletin of the American Astronomical Society*, 29:987, July 1997.

[254] G. C. Reid. Solar total irradiance variations and the global sea surface temperature record. *Journal of Geophys. Research*, 96:2835–2844, 1991.

[255] D. Rex, J. Bendisch, P. Eichler, and J. Zhang. Protecting and manoeuvring of spacecraft in space debris environment. *Advances in Space Research*, 11:53–62, 1991.

[256] E. Ribes, B. Beardsley, T. M. Brown, P. Delache, F. Laclare, J. R. Kuhn, and N. V. Leister. *The variability of the solar diameter*, pages 59–97. The Sun in Time, 1991.

[257] E. Ribes, J. C. Ribes, and R. Barthalot. Evidence for a larger sun with a slower rotation during the seventeenth century. *Nature*, 326:52–55, March 1987.

[258] A. J. Ridley, Y. Deng, and G. Tóth. The global ionosphere–thermosphere model. *Journal of Atmospheric and Terrestrial Physics*, 68:839–864, May 2006.

[259] M. Rieutord, T. Roudier, J. M. Malherbe, and F. Rincon. On mesogranulation, network formation and supergranulation. *Astronomy and Astrophysics*, 357:1063–1072, May 2000.

[260] E. Riklis, I. Emerit, and R. B. Setlow. New approaches to biochemical radioprotection: antioxidants and DNA repair enhancement. *Advances in Space Research*, 18:51–54, 1996.

[261] C. B. Roald, P. A. Sturrock, and R. Wolfson. Coronal Heating: Energy Release Associated with Chromospheric Magnetic Reconnection. *Astrophysical Journal*, 538:960–967, August 2000.

[262] T. Roudier, J. M. Malherbe, L. November, J. Vigneau, G. Coupinot, M. Lafon, and R. Muller. Intergranular plumes and formation of network bright points. *Astronomy and Astrophysics*, 320:605–611, April 1997.

[263] I. W. Roxburgh. Helioseismic Constraints on Solar Structure and the Solar Neutrino Problem. *Astrophysics and Space Science*, 261:57–58, 1998.

[264] M. A. Ruderman and J. W. Chamberlain. Origin of the sunspot modulation of ozone: Its implications for stratospheric NO injection. *Planetary and Space Science*, 23:247–268, February 1975.

[265] C. T. Russell. The dynamics of planetary magnetospheres. *Planetary and Space Science*, 49:1005–1030, August 2001.

[266] R. J. Rutten, R. H. Hammerschlag, F. M. Bettonvil, and P. Suetterlin. Dutch Open Telescope: Status and Prospects. *Bulletin of the American Astronomical Society*, 32:1290, October 2000.

[267] M. P. Ryutova and T. D. Tarbell. Observation of Shocks in the Chromosphere and Transition Region. *Bulletin of the American Astronomical Society*, 32:808, May 2000.

[268] M. Sanchez, F. Parra, M. Soler, and R. Soto. L'astrolabe DU ROA. Observations DU Soleil EN 1992. Observations of the Sun at the ROA astrolabe in 1992. *Astronomy and Astrophysics Supplement*, 110:351, April 1995.

[269] N. Sánchez-Ortiz, M. Belló-Mora, and H. Klinkrad. Collision Avoidance Manoeuvres during Spacecraft Mission Lifetime: Risk Reduction and Required DV. In *35th COSPAR Scientific Assembly*, page 2491, 2004.

[270] V. S. Sazonov and M. V. Yakovlev. On the explosive prevention of collisions of asteroidal and cometary bodies with the earth at their late detection. *Solar System Research*, 40:68–78, January 2006.

[271] N. Scafetta and B. J. West. Climate sensitivity to solar activity: The contribution of solar cycles 21-23 to global mean surface warming. *AGU Fall Meeting Abstracts*, December 2004.

[272] K. Schatten. Fair space weather for solar cycle 24. *Geophysical Research Letter*, 32:21106, November 2005.

[273] W. Schimmerling and F. A. Cucinotta. Critical radiation research for the Moon and Mars. In *35th COSPAR Scientific Assembly*, page 4325, 2004.

[274] C. J. Schrijver and C. Zwaan. *Solar and stellar magnetic activity.* Solar and stellar magnetic activity / Carolus J. Schrijver, Cornelius Zwaan. New York: Cambridge University Press, 2000. (Cambridge astrophysics series ; 34), 2000.

[275] D. W. Schwartzmann and T. Volk. *Nature*, 340:457–460, 1981.

[276] H. Sdunnus, J. Bendisch, and H. Klinkrad. The ESA MASTER'99 Space Debris and Meteoroid Reference Model. In H. Sawaya-Lacoste, editor, *ESA SP-473: Space Debris*, pages 299–307, October 2001.

[277] G. Seckmeyer and R. L. McKenzie. *Nature*, 359:135, 1992.

[278] S. Sello. Solar cycle forecasting: A nonlinear dynamics approach. *Astronomy and Astrophysics*, 377:312–320, October 2001.

[279] P. M. Sforza and J. L. Remo. NEO Mission Dynamics and Advanced Space Propulsion. In *Near-Earth Objects*, page 432, January 1997.

[280] M. R. Shavers, N. Zapp, R. E. Barber, J. W. Wilson, G. Qualls, L. Toupes, S. Ramsey, V. Vinci, G. Smith, and F. A. Cucinotta. Implementation of ALARA radiation protection on the ISS through polyethylene shielding augmentation of the Service Module Crew Quarters. *Advances in Space Research*, 34:1333–1337, 2004.

[281] N. J. Shaviv. On climate response to changes in the cosmic ray flux and radiative budget. *Journal of Geophysical Research (Space Physics)*, 110:8105, August 2005.

[282] H. Shibahashi. *Solar structure and the neutrino problem*, pages 231–246. Stellar astrophysical fluid dynamics. Edited by Michael J. Thompson, Jørgen Christensen-Dalsgaard. Cambridge (UK): Cambridge University Press, ISBN 0-521-81809-5, 2003, p. 231 - 246, 2003.

[283] H. Shinagawa. Our Current Understanding of the Ionosphere of Mars. *Advances in Space Research*, 26:1599–1608, 2000.

[284] E. M. Shoemaker and C. S. Shoemaker. *The Collision of Solid Bodies*, page 259. The New Solar System, 1990.

[285] J. A. Simpson. A Brief History of Recurrent Solar Modulation of the Galactic Cosmic Rays (1937-1990). *Space Science Reviews*, 83:169–176, January 1998.

[286] K. R. Sivaraman, S. S. Gupta, W. C. Livingston, L. Damé, W. Kalkofen, C. U. Keller, R. Smartt, and S. S. Hasan. Results from a revisit to the K_{2V} bright points. *Astronomy and Astrophysics*, 363:279–288, November 2000.

[287] R. Skartlien, R. F. Stein, and Å. Nordlund. Excitation of Chromospheric Wave Transients by Collapsing Granules. *Astrophysical Journal*, 541:468–488, September 2000.

[288] A. Skumanich. Time Scales for CA II Emission Decay, Rotational Braking, and Lithium Depletion. *Astrophysical Journal*, 171:565, February 1972.

[289] H. B. Snodgrass. Torsional Oscillations: Vorticity; Solar Cycle Predictions. *Bulletin of the American Astronomical Society*, 33:893, May 2001.

[290] M. Sobotka. Fine Structures in Sunspots. In *ASSL Vol. 239: Motions in the Solar Atmosphere*, pages 71–97, 1999.

[291] S. Sofia and A. S. Endal. Nature of the climatically significant changes of the solar constant. *LPI Contributions*, 390:96, 1979.

[292] S. K. Solanki. Empirical Modelling and Thermal Structure of Sunspots. In *ASP Conf. Ser. 118: 1st Advances in Solar Physics Euroconference. Advances in Physics of Sunspots*, page 178, 1997.

[293] S. K. Solanki and M. Fligge. Reconstruction of Past Solar Irradiance. *Space Science Reviews*, 94:127–138, November 2000.

[294] C. P. Sonett. Sunspot time series - Spectrum from square law modulation of the Hale cycle. *Geophysical Research Letter*, 9:1313–1316, December 1982.

[295] C. P. Sonett and G. E. Williams. Solar periodicities expressed in varves from glacial Skilak Lake, southern Alaska. *Journal of Geophys. Research*, 90:12019, December 1985.

[296] Q. Song and W. Cao. In *Astrophys. in Asia and its Future, Ed. P.S. Chen, Yunan Obs. Chin. Acadm. Sc.*, page 139, 1999.

[297] E. A. Spiegel and J.-P. Zahn. The solar tachocline. *Astronomy and Astrophysics*, 265:106–114, November 1992.

[298] H. C. Spruit. Effect of spots on a star's radius and luminosity. *Astronomy and Astrophysics*, 108:348–355, April 1982.

[299] B. Stauffer. Long Term Climate Records from Polar Ice. *Space Science Reviews*, 94:321–336, November 2000.

[300] O. Steiner and A. Ferriz-Mas. Connecting solar radiance variability to the solar dynamo with the virial theorem. *Astronomische Nachrichten*, 326:190–193, 2005.

[301] E. C. Stone, A. C. Cummings, F. B. McDonald, B. C. Heikkila, N. Lal, and W. R. Webber. Voyager 1 Explores the Termination Shock Region and the Heliosheath Beyond. *Science*, 309:2017–2020, September 2005.

[302] L. H. Strous, P. R. Goode, and T. R. Rimmele. The Dynamics of the Excitation of Solar Oscillations. *Astrophysical Journal*, 535:1000–1013, June 2000.

[303] Y. Suzuki. Solar Neutrinos. *Space Science Reviews*, 85:91–104, August 1998.

[304] B. Svenonius and E. Olausson. *Paleogeol. Paleoclimate*, 26:87–97, 1979.

[305] H. Svensmark. Influence of Cosmic Rays on Earth's Climate. *Physical Review Letters*, 81:5027–5030, November 1998.

[306] H. Svensmark and E. Friis-Christensen. Variation of cosmic ray flux and global cloud coverage-a missing link in solar-climate relationships. *Journal of Atmospheric and Terrestrial Physics*, 59:1225–1232, July 1997.

[307] H. Y. Tada, J. R. Carter, Jr., B. E. Anspaugh, and R. G. Downing. In *The Solar Cell Radiation Handbook, 3rd Edition, NASA/JPL Publ.*, 1982.

[308] E. Tajika. Physical and Geochemical Conditions Required for the Initiation of Snowball Earth. In *Lunar and Planetary Institute Conference Abstracts*, page 1573, March 2001.

[309] M. Takata. In *Frontier of Neutrino Astrophysics*, 1993.

[310] M. Takata and H. Shibahashi. Solar Models Based on Helioseismology and the Solar Neutrino Problem. *Astrophysical Journal*, 504:1035, September 1998.

[311] Q. Tang, B. J. Pang, and W. Zhang. Collision Risk Assessment for a Spacecraft in the Space Debris Environment. In D. Danesy, editor, *ESA SP-589: 4th European Conference on Space Debris*, page 721, August 2005.

[312] I. Thomson. *Mutat. Res.*, 430, 1999.

[313] Y. Totsuka. KAMIOKANDE and Super-Kainiokande. In *The Sun and Beyond*, page 221, 1996.

[314] L. Townsend. Space Radiation Hazards on Human Missions to the Moon and Mars. *AGU Fall Meeting Abstracts*, pages C1+, December 2004.

[315] L. W. Townsend and J. W. Wilson. The Interplanetary Radiation Environment and Methods to Shield from it. In C. R. Stoker and C. Emmart, editors, *Strategies for Mars: A Guide to Human Exploration*, page 283, 1996.

[316] D. Tripathi, V. Bothmer, S. K. Solanki, R. Schwenn, M. Mierla, and G. Stenborg. Plasma dynamics of a prominence associated coronal mass ejection. In *IAU Symposium*, pages 401–402, November 2004.

[317] S. Turck-Chièze, S. Couvidat, A. G. Kosovichev, A. H. Gabriel, G. Berthomieu, A. S. Brun, J. Christensen-Dalsgaard, R. A. García, D. O. Gough, J. Provost, T. Roca-Cortes, I. W. Roxburgh, and R. K. Ulrich. Solar Neutrino Emission Deduced from a Seismic Model. *Astrophysical Journal Letters*, 555:L69–L73, July 2001.

[318] S. Turck-Chièze, R. A. García, S. Couvidat, R. K. Ulrich, L. Bertello, F. Varadi, A. G. Kosovichev, A. H. Gabriel, G. Berthomieu, A. S. Brun, I. Lopes, P. Pallé, J. Provost, J. M. Robillot, and T. Roca Cortés. Looking for Gravity-Mode Multiplets with the GOLF Experiment aboard SOHO. *Astrophysical Journal*, 604:455–468, March 2004.

[319] S. Turck-Chieze and I. Lopes. Toward a unified classical model of the sun - On the sensitivity of neutrinos and helioseismology to the microscopic physics. *Astrophysical Journal*, 408:347–367, May 1993.

[320] P. Ulmschneider, E. R. Priest, and R. Rosner, editors. *Mechanisms of Chromospheric and Coronal Heating*, 1991.

[321] R. K. Ulrich and L. Bertello. Solar-Cycle Dependence of the Sun's Apparent Radius in the Neutral Iron Spectral Line at 525-NM. *Nature*, 377:214, September 1995.

[322] I. G. Usoskin, S. V. Berdyugina, and J. Poutanen. Preferred sunspot longitudes: non-axisymmetry and differential rotation. *Astronomy and Astrophysics*, 441:347–352, October 2005.

[323] G. B. Valsecchi, A. Milani, A. Rossi, and G. Tommei. The SRT, Near-Earth objects, and space debris. *Memorie della Societa Astronomica Italiana Supplement*, 10:186, 2006.

[324] L. van Driel-Gesztelyi. CMEs and Magnetic Helicity. In *Solar Magnetic Phenomena, Proceedings of the 3rd Summerschool and Workshop held at the Solar Observatory Kanzelhöhe, Kärnten, Austria, August 25 – September 5, 2003. Edited by A. Hanslmeier, A. Veronig, and M. Messerotti. Astronomy and Astrophysics Space Science Library, vol. 320, Springer, Dordrecht, The Netherlands, 2005., p.115-138*, pages 57–85, 2005.

[325] M. Vazquez and A. Hanslmeier. *Ultraviolet radiation in the solar system.* Ultraviolet radiation in the solar system, by M. Vazquez and A. Hanslmeier. Astrophysics and space science library (ASSL), vol. 331. Dordrecht: Springer, 2006., 2006.

[326] A. Veronig, B. Vršnak, B. R. Dennis, M. Temmer, A. Hanslmeier, and J. Magdalenić. Investigation of the Neupert effect in solar flares. I. Statistical properties and the evaporation model. *Astronomy and Astrophysics*, 392:699–712, September 2002.

[327] B. Vršnak. Solar flares - observations and theory. In *Solar Magnetic Phenomena, Proceedings of the 3rd Summerschool and Workshop held at the Solar Observatory Kanzelhöhe, Kärnten, Austria, August 25 – September 5, 2003. Edited by A. Hanslmeier, A. Veronig, and M. Messerotti. Astronomy and Astrophysics Space Science Library, vol. 320, Springer, Dordrecht, The Netherlands, 2005., p.27-56*, pages 27–56, 2005.

[328] J. C. G. Walker, P. B. Hays, and J. F. Kasting. A negative feedback mechanism for the long-term stabilization of the earth's surface temperature. *Journal of Geophys. Research*, 86:9776–9782, October 1981.

[329] R. W. Walsh. Solar Magnetohydrodynamics. In *The Dynamic Sun, Proceedings of the Summerschool and Workshop held at the Solar Observatory*, page 129, May 2001.

[330] G. Walther. On the Solar-Cycle Modulation of the Homestake Solar Neutrino Capture Rate and the Shuffle Test. *Astrophysical Journal*, 513:990–996, March 1999.

[331] S. Watanabe and H. Shibahashi. Solar Models with Helioseismic Constraints and the Solar Neutrino Problem. *Publications of the Astronomical Society of Japan*, 53:565–575, June 2001.

[332] D. F. Webb. CMEs and the solar cycle variation in their geoeffectiveness. In *ESA SP-508: From Solar Min to Max: Half a Solar Cycle with SOHO*, pages 409–419, June 2002.

[333] R. Wieler and H. Baur. Fractionation of Xe, Kr, and AR in the Solar Corpuscular Radiation Deduced by Closed System Etching of Lunar Soils. *Astrophysical Journal*, 453:987, November 1995.

[334] R. Wieler, P. Etique, P. Signer, and G. Poupeau. Record of the solar corpuscular radiation in minerals from lunar soils - A comparative study of noble gases and tracks. In *Lunar and Planetary Science Conference*, pages 1369–1393, 1980.

[335] L. A. Willson, G. H. Bowen, and C. Struck-Marcell. Mass loss on the main sequence. *Comments on Astrophysics*, 12:17–34, April 1987.

[336] R. C. Willson. Solar total irradiance observations by active cavity radiometers. *Solar Physics*, 74:217–229, November 1981.

[337] R. C. Willson. Measurements of solar total irradiance and its variability. *Space Science Reviews*, 38:203–242, August 1984.

[338] R. C. Willson and H. S. Hudson. Solar luminosity variations in solar cycle 21. *Nature*, 332:810–812, April 1988.

[339] J. W. Wilson, R. K. Tripathi, G. D. Qualls, F. A. Cucinotta, R. E. Prael, J. W. Norbury, J. H. Heinbockel, and J. Tweed. A space radiation transport method development. *Advances in Space Research*, 34:1319–1327, 2004.

[340] A. D. Wittmann, E. Alge, and M. Bianda. Detection of a significant change in the solar diameter. *Solar Physics*, 145:205, May 1993.

[341] J.-G. Wu, L. Eliasson, H. Lundstedt, A. Hilgers, L. Andersson, and O. Norberg. Space Environment Effects on Geostationary Spacecraft: Analysis and Prediction. *Advances in Space Research*, 26:31–36, 2000.

[342] T. C. Yang, M. Mei, K. A. George, and L. M. Craise. DNA damage and repair in oncogenic transformation by heavy ion radiation. *Advances in Space Research*, 18:149–158, 1996.

[343] C.-T. Yeh, M. D. Ding, and P. F. Chen. Waiting time distribution of CMEs. In *Solar Magnetic Phenomena, Proceedings of the 3rd Summerschool and Workshop held at the Solar Observatory Kanzelhöhe, Kärnten, Austria, August 25 – September 5, 2003. Edited by A. Hanslmeier, A. Veronig, and M. Messerotti. Astronomy and Astrophysics Space Science Library, vol. 320, Springer, Dordrecht, The Netherlands, 2005., p.171-174*, pages 171–174, 2005.

[344] K. J. Zahnle and J. C. G. Walker. The evolution of solar ultraviolet luminosity. *Reviews of Geophysics and Space Physics*, 20:280–292, May 1982.

[345] E. J. Zeller and L. B. Ronca. Space Weathering of Lunar and Asteroidal Surfaces. *Icarus*, 7:372–379, 1967.

[346] J. F. Ziegler and W. A. Lanford. Effect of Cosmic Rays on Computer Memories. *Science*, 206:776–788, November 1979.

[347] R. D. Zwickl, K. A. Doggett, S. Sahm, W. P. Barrett, R. N. Grubb, T. R. Detman, V. J. Raben, C. W. Smith, P. Riley, R. E. Gold, R. A. Mewaldt, and T. Maruyama. The NOAA Real-Time Solar-Wind (RTSW) System using ACE Data. *Space Science Reviews*, 86:633–648, 1998.

Internet

Today's space weather can be found under:

http://www.sec.noaa.gov/today.html and http://www.windows.ucar.edu/spaceweather/

NASA's space environment center (SEC):

http://www.sec.noaa.gov/

The web site of the National Oceanic and Atmospheric Administration:

http://www.sec.noaa.gov/

ESA Space Weather Site:

http://www.estec.esa.nl/wmwww/spweather/

NASA Space weather resources:

http://spdf.gsfc.nasa.gov/space_weather/Space_Weather_at_SSDOO.html

Space Science Institute/ NASA and NSF site:

http://www.spacescience.org/SWOP/

Lund Space Weather Center

http://www.irfl.lu.se/

Australian Space Weather

http://www.ips.gov.au/

Further references can be found in these sites.

List of Tables

Index

Astrophysics and Space Science Library

Volume 347: ***The Sun and Space Weather***, by A.Hanslmeier.
Hardbound ISBN 1-4020-5603-6, February 2007

Volume 346: ***Exploring the Secrets of the Aurora***, by S.-I. Akasofu.
Hardbound ISBN 0-387-45094-7, April 2007

Volume 345: ***Canonical Perturbation Theories: Degenerate Systems and Resonance***, by S. Ferraz-Mello.
Hardbound ISBN 0-387-38900-8, December 2006

Volume 344: ***Space Weather: Research towards Applications in Europe***, edited by J. Lilensten.
Hardbound ISBN 1-4020-5445-9, January 2007

Volume 343: ***Organizations and Strategies in Astronomy: Volume 7***, edited by A. Heck.
Hardbound ISBN 1-4020-5300-2, December 2006

Volume 342: ***The Astrophysics of Emission Line Stars***, by T. Kogure and K.-C. Leung.
Hardbound ISBN 0-387-34500-0, February 2007

Volume 341: ***Plasma Astrophysics, Part II: Reconnection and Flares***, by B.V. Somov.
Hardbound ISBN 0-387-34948-0, November 2006

Volume 340: ***Plasma Astrophysics, Part I: Fundamentals and Practice***, by B.V. Somov.
Hardbound ISBN 0-387-34916-2, 2006

Volume 339: ***Cosmic Ray Interactions, Propagation, and Acceleration in Space Plasmas***, edited by L. Dorman.
Hardbound ISBN 978-1-4020-5100-5, September 2006

Volume 338: ***Solar Journey: The Significance of Our Galactic Environment for the Heliospere and Earth***, edited by P. Frisch.
Hardbound ISBN 978-1-4020-4397-0, September 2006

Volume 337: ***Astrophysical Disks,*** edited by A. M. Fridman, M. Y. Marov, I.G. Kovalenko.
Hardbound ISBN 1-4020-4347-3, June 2006

Volume 336: ***Scientific Detectors for Astronomy 2005***, edited by J.E. Beletic, J.W. Beletic, P. Amico.
Hardbound ISBN 1-4020-4329-5, December 2005

Volume 335: ***Organizations and Strategies in Astronomy 6***, edited by A. Heck.
Hardbound ISBN 1-4020-4055-5, November 2005

Volume 334: ***The New Astronomy: Opening the Electromagnetic Window and Expanding our View of Planet Earth***, edited by W. Orchiston.
Hardbound ISBN 1-4020-3723-6, October 2005

Volume 333: ***Planet Mercury***, by P. Clark and S. McKenna-Lawlor.
Hardbound ISBN 0-387-26358-6, November 2005

Volume 332: ***White Dwarfs: Cosmological and Galactic Probes,*** edited by E.M. Sion, S. Vennes, H.L. Shipman.
Hardbound ISBN 1-4020-3693-0, September 2005

Volume 331: ***Ultraviolet Radiation in the Solar System***, by M. Vázquez and A. Hanslmeier.
Hardbound ISBN 1-4020-3726-0, November 2005

Volume 330: ***The Multinational History of Strasbourg Astronomical Observatory,*** edited by A. Heck.
Hardbound ISBN 1-4020-3643-4, June 2005

Volume 329: ***Starbursts – From 30 Doradus to Lyman Break Galaxies,*** edited by R. de Grijs, R.M. González Delgado.
Hardbound ISBN 1-4020-3538-1, May 2005

Volume 328: ***Comets***, by J.A. Fernández.
Hardbound ISBN 1-4020-3490-3, July 2005

Volume 327: ***The Initial Mass Function 50 Years Later***, edited by E. Corbelli, F. Palla, H. Zinnecker.
Hardbound ISBN 1-4020-3406-7, June 2005

Volume 325: ***Kristian Birkeland – The First Space Scientist,*** by A. Egeland, W.J. Burke.
Hardbound ISBN 1-4020-3293-5, April 2005

Volume 324: ***Cores to Clusters – Star Formation with next Generation Telescopes,*** edited by M.S. Nanda Kumar, M. Tafalla, P. Caselli.
Hardbound ISBN 0-387-26322-5, October 2005

Volume 323: ***Recollections of Tucson Operations,*** by M.A. Gordon.
Hardbound ISBN 1-4020-3235-8, December 2004

Volume 322: ***Light Pollution Handbook,*** by K. Narisada, D. Schreuder
Hardbound ISBN 1-4020-2665-X, November 2004

Volume 321: ***Nonequilibrium Phenomena in Plasmas***, edited by A.S. Shrama, P.K. Kaw.
Hardbound ISBN 1-4020-3108-4, December 2004

Volume 320: ***Solar Magnetic Phenomena,*** edited by A. Hanslmeier, A. Veronig, M. Messerotti.
Hardbound ISBN 1-4020-2961-6, December 2004

Volume 319: ***Penetrating Bars through Masks of Cosmic Dust***, edited by D.L. Block, I. Puerari, K.C. Freeman, R. Groess, E.K. Block.
Hardbound ISBN 1-4020-2861-X, December 2004

Volume 318: ***Transfer of Polarized light in Planetary Atmospheres***, by J.W. Hovenier, C. van der Mee, J.W. Domke.
Hardbound ISBN 1-4020-2855-5.
Softcover ISBN 1-4020-2889-X, November 2004

Volume 317: ***The Sun and the Heliosphere as an Integrated System,*** edited by G. Poletto, S.T. Suess.
Hardbound ISBN 1-4020-2830-X, November 2004

Volume 316: *Civic Astronomy - Albany's Dudley Observatory, 1852-2002,* by G. Wise.
Hardbound ISBN 1-4020-2677-3, October 2004

Volume 315: ***How does the Galaxy Work*** *- A Galactic Tertulia with Don Cox and* ***Ron Reynolds,*** edited by E. J. Alfaro, E. Pérez, J. Franco
Hardbound ISBN 1-4020-2619-6, September 2004

Volume 314: ***Solar and Space Weather Radiophysics- Current Status and Future Developments,*** edited by D.E. Gary and C.U. Keller
Hardbound ISBN 1-4020-2813-X, August 2004

Volume 313: ***Adventures in Order and Chaos***, by G. Contopoulos.
Hardbound ISBN 1-4020-3039-8, January 2005

Volume 312: ***High-Velocity Clouds,*** edited by H. van Woerden, U. Schwarz, B. Wakker.
Hardbound ISBN 1-4020-2813-X, September 2004

Volume 311: ***The New ROSETTA Targets- Observations, Simulations and Instrument Performances***, edited by L. Colangeli, E. Mazzotta Epifani, P. Palumbo.
Hardbound ISBN 1-4020-2572-6, September 2004

Volume 310: ***Organizations and Strategies in Astronomy 5,*** edited by A. Heck
Hardbound ISBN 1-4020-2570-X, September 2004

Volume 309: ***Soft X-ray Emission from Clusters of Galaxies and Related Phenomena***, edited by R. Lieu and J. Mittaz
Hardbound ISBN 1-4020-2563-7, September 2004

Volume 308: ***Supermassive Black Holes in the Distant Universe,*** edited by A.J. Barger.
Hardbound ISBN 1-4020-2470-3, August 2004

Volume 307: ***Polarization in Spectral Lines***, by E. Landi Degl'Innocenti and M. Landolfi.
Hardbound ISBN 1-4020-2414-2, August 2004

Volume 306: ***Polytropes – Applications in Astrophysics and Related Fields,*** by G.P. Horedt.
Hardbound ISBN 1-4020-2350-2, September 2004

Volume 305: ***Astrobiology: Future Perspectives,*** edited by P. Ehrenfreund, W.M. Irvine, T. Owen, L. Becker, J. Blank, J.R. Brucato, L. Colangeli, S. Derenne, A. Dutrey, D. Despois, A. Lazcano, F. Robert
Hardbound ISBN 1-4020-2304-9, July 2004
Paperback ISBN 1-4020-2587-4, July 2004

Volume 304: ***Cosmic Gammy-ray Sources,*** edited by K.S. Cheng and G.E. Romero
Hardbound ISBN 1-4020-2255-7, September 2004

Volume 303: ***Cosmic rays in the Earth's Atmosphere and Underground***, by L.I, Dorman.
Hardbound ISBN 1-4020-2071-6, August 2004

Volume 302:***Stellar Collapse,*** edited by Chris L. Fryer
Hardbound, ISBN 1-4020-1992-0, April 2004

Volume 301: ***Multiwavelength Cosmology***, edited by Manolis Plionis
Hardbound, ISBN 1-4020-1971-8, March 2004

Volume 300:***Scientific Detectors for Astronomy,*** edited by Paola Amico, James W. Beletic, Jenna E. Beletic
Hardbound, ISBN 1-4020-1788-X, February 2004

Volume 299: ***Open Issues in Local Star Fomation,*** edited by Jacques Lépine, Jane Gregorio-Hetem.
Hardbound, ISBN 1-4020-1755-3, December 2003

Volume 298: ***Stellar Astrophysics - A Tribute to Helmut A. Abt,*** edited by K.S. Cheng, Kam Ching Leung, T.P. Li.
Hardbound, ISBN 1-4020-1683-2, November 2003

Volume 297: ***Radiation Hazard in Space,*** by Leonty I. Miroshnichenko
Hardbound, ISBN 1-4020-1538-0, September 2003

Volume 296: ***Organizations and Strategies in Astronomy, volume 4,*** edited by André Heck.
Hardbound, ISBN 1-4020-1526-7, October 2003

Volume 295: ***Integrable Problems of Celestial Mechanics in Spaces of Constant Curvature,*** by T.G. Vozmischeva.
Hardbound, ISBN 1-4020-1521-6, October 2003

Volume 294: ***An Introduction to Plasma Astrophysics and Magnetohydrodynamics,*** by Marcel Goossens
Hardbound, ISBN 1-4020-1429-5, August 2003
Paperback, ISBN 1-4020-1433-3, August 2003

Volume 293: ***Physics of the Solar System,*** by Bruno Bertotti, Paolo Farinella, David Vokrouhlický
Hardbound, ISBN 1-4020-1428-7, August 2003
Paperback, ISBN 1-4020-1509-7, August 2003

Volume 292: ***Whatever Shines Should Be Observed,*** by Susan M.P. McKenna-Lawlor.
Hardbound, ISBN 1-4020-1424-4, September 2003

Volume 291: ***Dynamical Systems and Cosmology***, by Alan Coley
Hardbound, ISBN 1-4020-1403-1, November 2003

Volume 290: ***Astronomy Communication,*** edited by André Heck, Claus Madsen
Hardbound, ISBN 1-4020-1345-0, July 2003

Volume 287/8/9: ***The Future of Small Telescopes in the New Millennium,*** edited by Terry D. Oswalt.
Hardbound Set only of 3 volumes, ISBN 1-4020-0951-8, July 2003

Volume 286: ***Searching the Heavens and the Earth: The History of Jesuit Observatories,*** by Agustín Udías
Hardbound, ISBN 1-4020-1189-X, October 2003

Volume 285: ***Information Handling in Astronomy - Historical Vistas,*** edited by André Heck
Hardbound, ISBN 1-4020-1178-4, March 2003

Volume 284: ***Light Pollution: The Global View,*** edited by Hugo E. Schwarz
Hardbound, ISBN 1-4020-1174-1, April 2003

Volume 283: ***Mass-Losing Pulsating Stars and Their Circumstellar Matter,*** edited by Y. Nakada, M. Honma, M. Seki
Hardbound, ISBN 1-4020-1162-8, March 2003

Volume 282: ***Radio Recombination Lines,*** by M.A. Gordon, R.L. Sorochenko
Hardbound, ISBN 1-4020-1016-8, November 2002

Volume 281: ***The IGM/Galaxy Connection,*** edited by Jessica L. Rosenberg, Mary E. Putman
Hardbound, ISBN 1-4020-1289-6, April 2003

Volume 280: ***Organizations and Strategies in Astronomy III,*** edited by André Heck
Hardbound, ISBN 1-4020-0812-0, September 2002

Volume 279: ***Plasma Astrophysics , Second Edition,*** by Arnold O. Benz
Hardbound, ISBN 1-4020-0695-0, July 2002

Volume 278: ***Exploring the Secrets of the Aurora,*** by Syun-Ichi Akasofu
Hardbound, ISBN 1-4020-0685-3, August 2002

Volume 277: ***The Sun and Space Weather,*** by Arnold Hanslmeier
Hardbound, ISBN 1-4020-0684-5, July 2002

Volume 276: ***Modern Theoretical and Observational Cosmology,*** edited by Manolis Plionis, Spiros Cotsakis
Hardbound, ISBN 1-4020-0808-2, September 2002

Volume 275: ***History of Oriental Astronomy,*** edited by S.M. Razaullah Ansari
Hardbound, ISBN 1-4020-0657-8, December 2002

Volume 274: ***New Quests in Stellar Astrophysics: The Link Between Stars and Cosmology,*** edited by Miguel Chávez, Alessandro Bressan, Alberto Buzzoni,Divakara Mayya
Hardbound, ISBN 1-4020-0644-6, June 2002

Volume 273: ***Lunar Gravimetry,*** by Rune Floberghagen
Hardbound, ISBN 1-4020-0544-X, May 2002

Volume 272:***Merging Processes in Galaxy Clusters,*** edited by L. Feretti, I.M. Gioia, G. Giovannini
Hardbound, ISBN 1-4020-0531-8, May 2002

Volume 271: ***Astronomy-inspired Atomic and Molecular Physics,*** by A.R.P. Rau
Hardbound, ISBN 1-4020-0467-2, March 2002

Volume 270: ***Dayside and Polar Cap Aurora,*** by Per Even Sandholt, Herbert C. Carlson, Alv Egeland
Hardbound, ISBN 1-4020-0447-8, July 2002

Volume 269: ***Mechanics of Turbulence of Multicomponent Gases,*** by Mikhail Ya. Marov, Aleksander V. Kolesnichenko
Hardbound, ISBN 1-4020-0103-7, December 2001

Volume 268: ***Multielement System Design in Astronomy and Radio Science,*** by Lazarus E. Kopilovich, Leonid G. Sodin
Hardbound, ISBN 1-4020-0069-3, November 2001

Volume 267: ***The Nature of Unidentified Galactic High-Energy Gamma-Ray Sources,*** edited by Alberto Carramiñana, Olaf Reimer, David J. Thompson
Hardbound, ISBN 1-4020-0010-3, October 2001

Volume 266: *Organizations and Strategies in Astronomy II*,
edited by André Heck
Hardbound, ISBN 0-7923-7172-0, October 2001

Volume 265: *Post-AGB Objects as a Phase of Stellar Evolution*,
edited by R. Szczerba, S.K. Górny
Hardbound, ISBN 0-7923-7145-3, July 2001

Volume 264: *The Influence of Binaries on Stellar Population Studies*,
edited by Dany Vanbeveren
Hardbound, ISBN 0-7923-7104-6, July 2001

Volume 262: *Whistler Phenomena - Short Impulse Propagation*,
by Csaba Ferencz, Orsolya E. Ferencz, Dániel Hamar, János Lichtenberger
Hardbound, ISBN 0-7923-6995-5, June 2001

Volume 261: *Collisional Processes in the Solar System*,
edited by Mikhail Ya. Marov, Hans Rickman
Hardbound, ISBN 0-7923-6946-7, May 2001

Volume 260: *Solar Cosmic Rays*, by Leonty I. Miroshnichenko
Hardbound, ISBN 0-7923-6928-9, May 2001

For further information about this book series we refer you to the following web site:
www.springer.com

To contact the Publishing Editor for new book proposals:
Dr. Harry (J.J.) Blom: harry.blom@springer.com
Sonja Japenga: sonja.japenga@springer.com