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

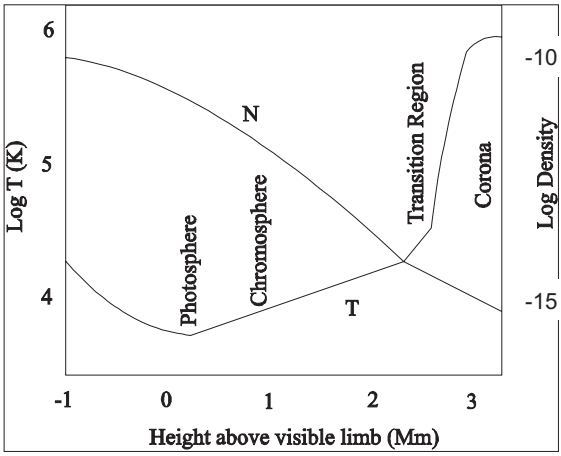

Figure 3.1: Variation of electron temperature and electron density in the solar atmosphere

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-

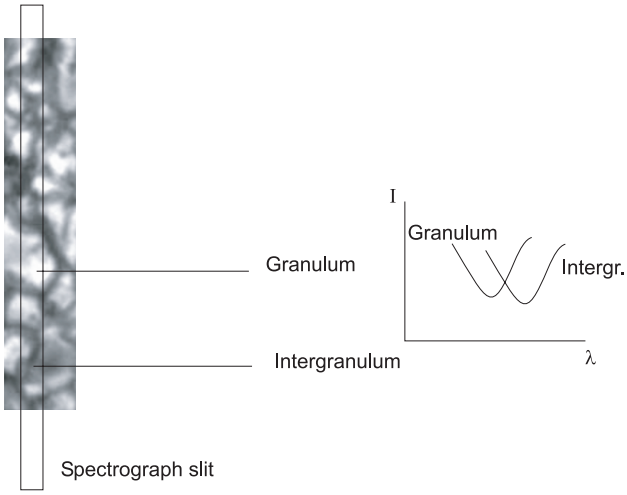

Figure 3.2: Spectroscopic observation of solar granulation. The entrance of a spectrograph slit covers different granular/intergranular areas. Line profiles emanating from granules are blueshifted because matter moves upwards and profiles from intergranular areas are redshifted because matter moves away 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

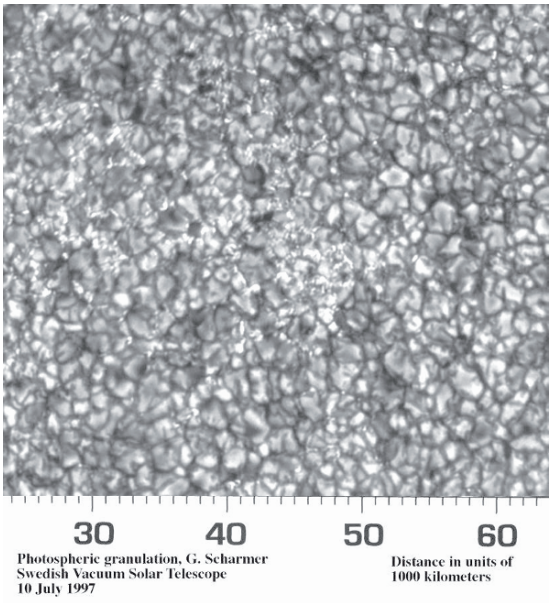

Figure 3.3: Solar granulation and small network bright points

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

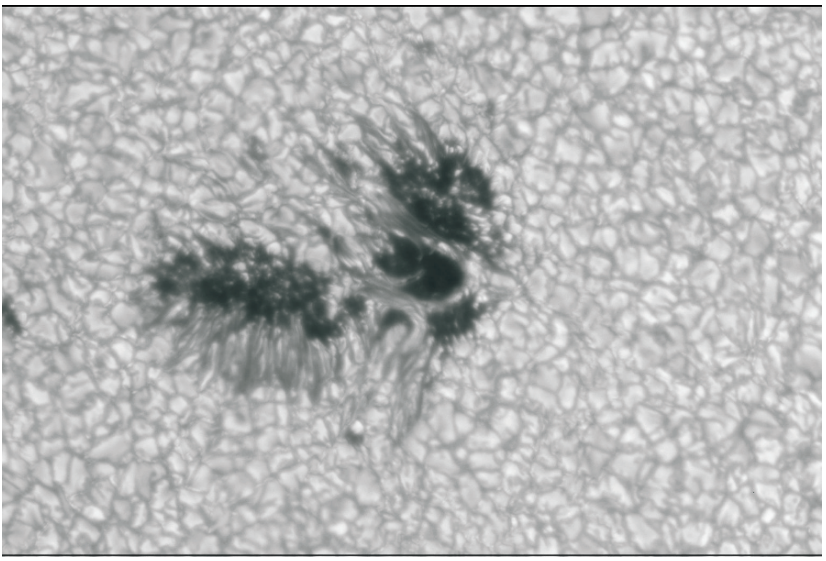

Figure 3.4: Large sunspot showing the dark central umbra and the filamentary penumbra. Outside the penumbra the granulation pattern is clearly seen. Courtesy: M. Sobotka, A.H., SST, La Palma, 2003

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° 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 $\leq$ 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.

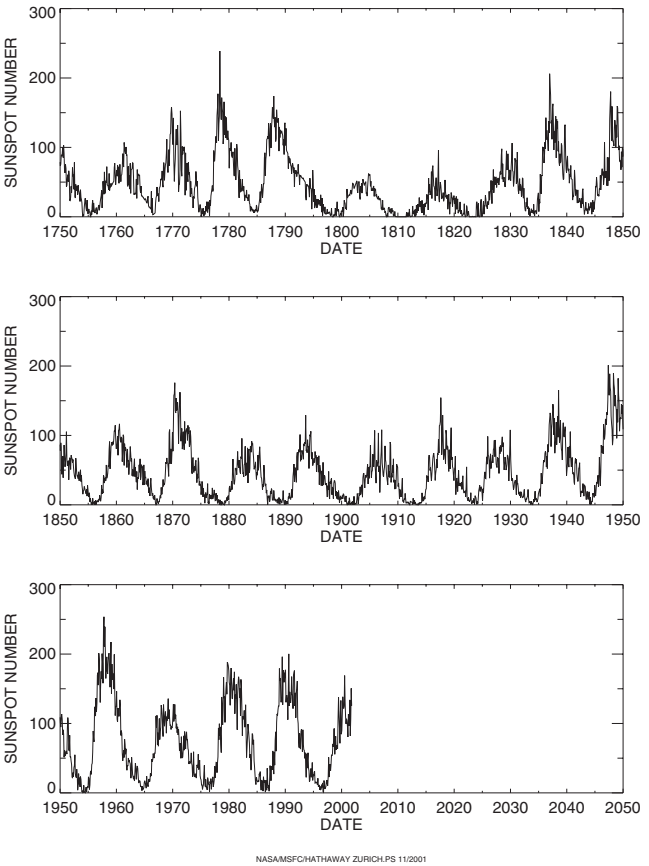

Figure 3.5: Relative Sunspot number. Two cycles can bee seen, the normal cycle with about 11 years and the so called Gleissberg cycle with a period of about 80 years.

- 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

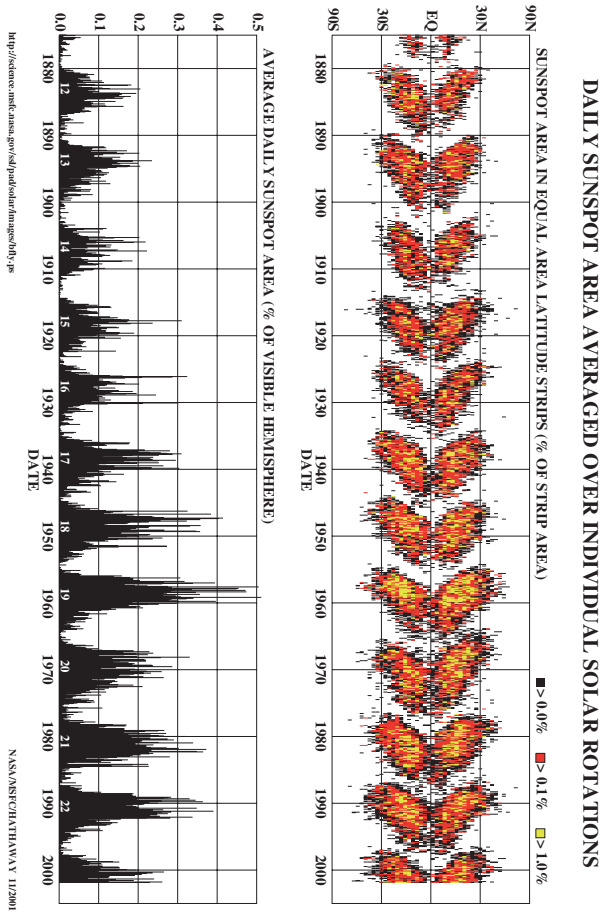

Figure 3.6: Butterflydiagram illustrating the equatorward motion of spots during the activity cycle.

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].

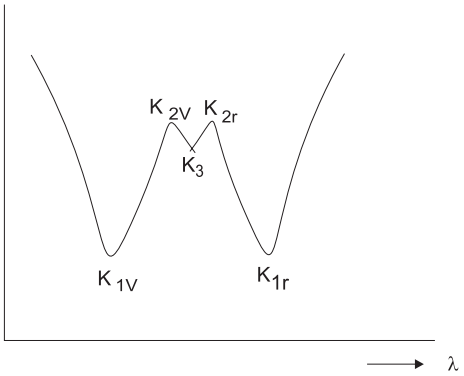

Figure 3.7: Profile of the CaII line

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

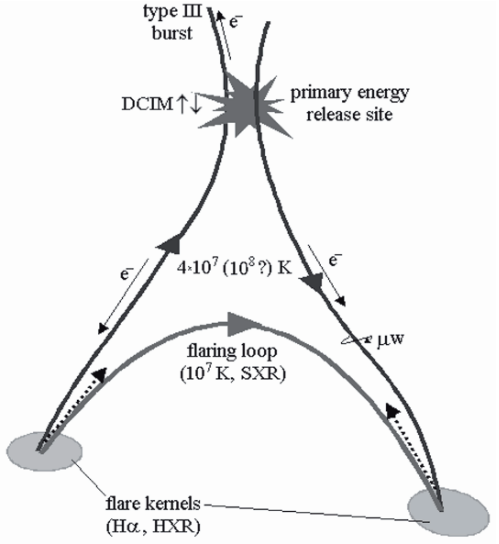

Figure 3.8: Summary of basic processes of most solar flares and their emission regions. Courtesy: B. Vrsnjak.

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	$< 1\,\mathrm{h}$	$> 1\,\mathrm{h}$
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 $10^{24}\,\mathrm{J}$ 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

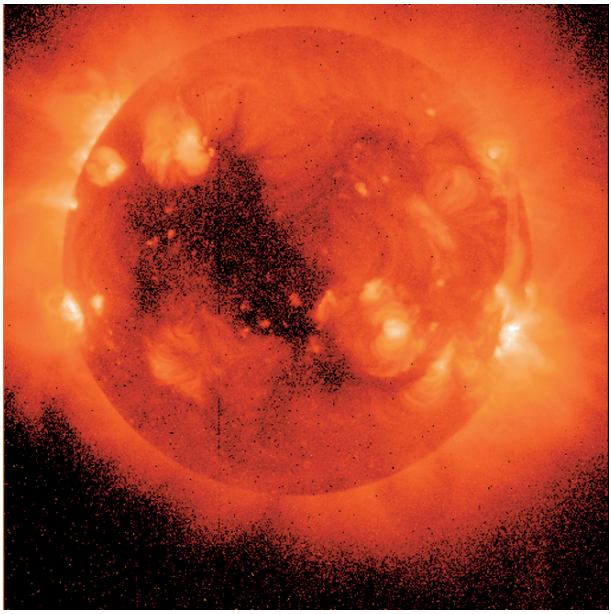

Figure 3.9: Coronal hole seen by the solar satellite YOHKOH

"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

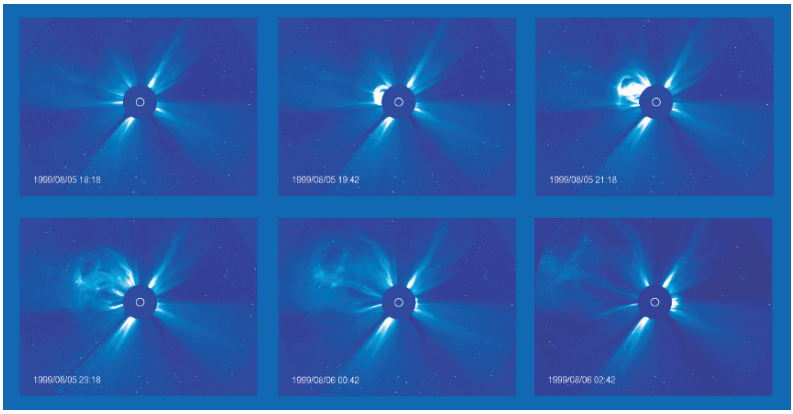

Figure 3.10: Progress of a Coronal Mass Ejection (CME) observed over an eight hour period on 5-6 August 1999 by LASCO C3. The dark disk blocks the Sun so that the LASCO instrument can observe the structures of the corona in visible light. The white circle represents the size and position of the Sun. 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 (≤ 800 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:

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

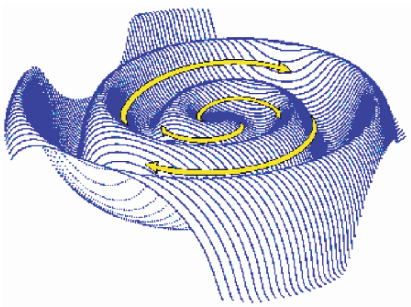

Figure 3.12: The Parker spiral. Courtesy: J. Jokipii, University of Arizona

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.

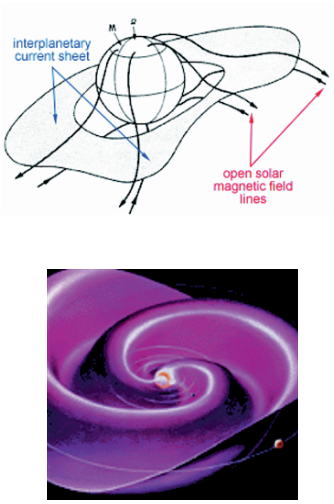

Figure 3.13: The interplanetary current sheet (above) and the heliospheric current sheet (below). Courtesy: http://pluto.space.swri.edu/image/glossary/IMF.html

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

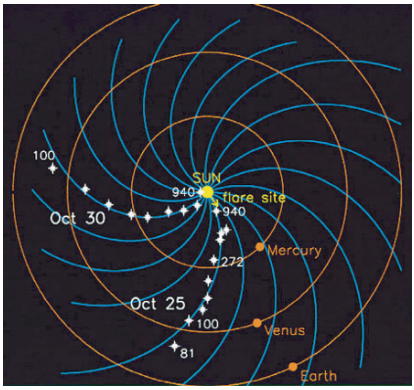

Figure 3.14: The spiral nature of the IMF. Courtesy: NASA

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).