45

SAMI K. SOLANKI* AND REINER HAMMER**

The solar atmosphere

As a typical star, and the only one that can be spatially resolved by direct means, the study of the Sun has provided an insight into many of the fundamental processes taking place in stellar atmospheres, often at small scales. A prime example is magneto-convection or the formation of coronae and the consequent emission of copious amounts of X-rays. In addition, the Sun's apparent brightness allows measurements with unprecedented accuracy. Thus the Sun is the standard against which cosmic abundances are compared. Its high apparent brightness also means that the Sun is a strong source at almost all wavelengths and thus detectable with simple, not particularly sensitive equipment such as the early instruments flown in space. Thus for many wavelengths the Sun was the first (or one of the first) cosmic source(s) detected.

However, only the lowest layers of the Sun's atmosphere, the photosphere and chromosphere, can be regularly observed from the ground over the solar disk. The transition region, corona and the solar wind are best studied from space, and even many properties of the photosphere (such as the variation of solar irradiance with time) had to await space-based observations for their determination or discovery.

1 OVERVIEW OF THE SOLAR ATMOSPHERE

Traditionally the atmosphere of the Sun is divided into four layers, starting with the photosphere at the bottom, moving up through the chromosphere and transition region to the corona. The photosphere is the layer in which the temperature drops outwards from around 5800 K at the solar surface to around 4000 K at the temperature minimum. Beyond that point it rises again, first relatively gently (forming the chromospheric plateau), but then very rapidly in the transition region (TR). The temperature profile becomes flatter again in the corona. The boundary between the corona and the TR is often drawn at approximately 10^6 K. This boundary, like that between chromosphere and TR, is not sharp or well defined. At still greater distances from the solar surface the temperature gradually decreases again, achieving values of approximately 10^5 K at 1 AU (whereby electrons and ions need not have the same temperature in the heliosphere). As we shall see in subsequent sections, the simple plane-parallel representation of the solar gas outlined above is not tenable in any layer of the atmosphere. At any given height more than one atmospheric component is present, each having its own temperature, density and velocity structure.

Features as diverse as granular convection cells in the photosphere (Figure 1) and magnetic loops in the corona (Figure 2) are now known to structure the respective layers of the atmosphere. In addition to being spatially inhomogeneous at almost all spatial scales, the solar atmosphere is also highly dynamic at almost all timescales. Much of the interesting physics to be learnt by studying the solar atmosphere is related to this structuring and dynamics and the associated heating of the chromosphere and corona.

In the following we discuss the various atmospheric layers, starting with the photosphere and moving outward. Particular emphasis is placed on the contributions made by space missions to our knowledge and understanding of the solar atmosphere. Since these contributions are largest for the transition region and corona our discussion of these layers will be more detailed than of the photosphere and chromosphere. Table 1 summarizes the space missions mentioned in this chapter.

2 THE PHOTOSPHERE

2.1 The plane-parallel photosphere

The solar photosphere is the layer that emits most of the solar radiative energy flux, with the emitted spectrum

* Max-Planck-Institut für Aeronomie, Kaflenburg-Lindau, Germany
** Kiepenheuer-Institut für Sonnenphysik, Freiburg i/B, Germany

Sami K. Solanki and Reiner Hammer, The Century of Space Science, 1–24

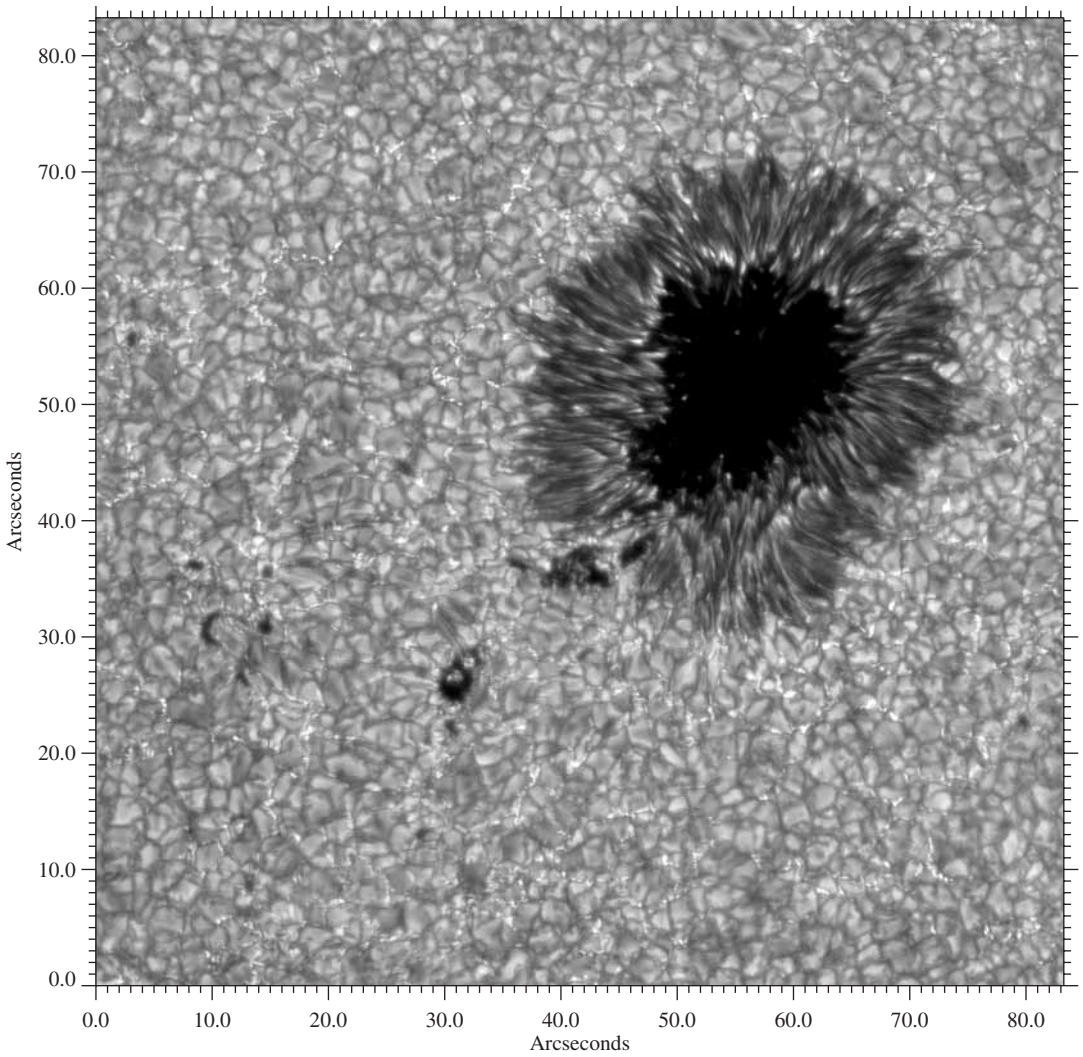

Figure 1 A snapshot of a part of the solar photosphere taken with a filter centred on the g band at 430.5 nm by T. Berger and G. Scharmer. The image covers 60 000 × 60 000 km on the Sun. The most prominent feature is a sunspot. The much smaller dark features are pores. Also visible are granules (bright cells surrounded by dark lanes) and bright points corresponding to magnetic elements. (Courtesy of T. Berger.)

having its peak in the visible (in the green part of the wavelength range). As such, the photosphere is the atmospheric layer most easily observed from the ground and consequently the one to whose investigation spacecraft have contributed the least. This, however, is changing at a rapid pace, with the ESA–NASA Solar and Heliospheric Observatory (SOHO; Fleck and Domingo 1995) providing the first glimpses of how space-based telescopes can revolutionize our understanding of the photosphere. The next major highlight is expected to be provided by the Japan–US–UK Solar B mission.

The brightness across the solar disk is not constant but rather decreases from the centre of the disk to its edge (the solar limb) at visible wavelengths. This is called limb darkening. Since at the limb the radiation is emitted at greater heights, limb darkening implies a decrease in the temperature with height. Furthermore, the spectral form of the limb darkening provides information on the continuum absorption coefficient. Such observations confirmed the proposal by Wildt (1939) that in the visible the absorption is dominated by the H^- ion in spite of its low abundance (Chalonge and Kourganoff 1946).

Traditionally the limb darkening and the shapes and strengths of absorption lines (Fraunhofer lines) have been employed to determine the temperature stratification in the solar photosphere. These diagnostics reveal that the temperature decreases outwards in the solar photosphere from over 6500 K at the deepest observable layers to around 4000 K at the temperature minimum (e.g. Holweger 1967). The advent of UV observations from space, in particular

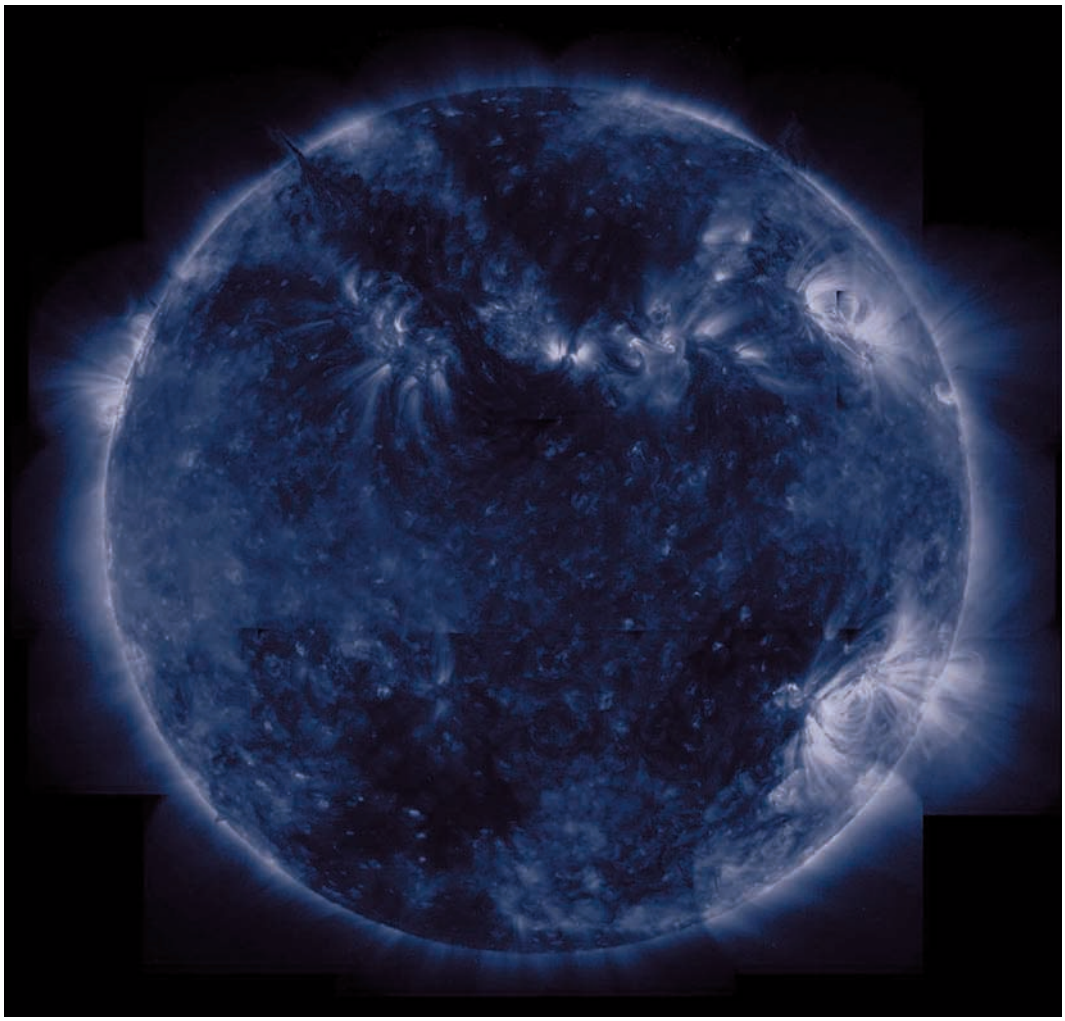

Figure 2 Composite of several high-resolution images taken with the Transition Region and Coronal Explorer (TRACE; Handy *et al*. 1999) in a spectral band near 171 Å, which is dominated by emission from eightfold ionized iron atoms (Fe IX) formed around 10^6 K. At these temperatures the network is no longer visible, and the disk emission is dominated by active regions, by coronal loops in the quiet corona (i.e. outside of coronal holes and active regions), and by numerous bright points. Plumes extend as ray-shaped density enhancements from the north and south polar coronal holes. (Courtesy of the TRACE team. TRACE is a mission of the Stanford–Lockheed Institute for Space Research, and part of the NASA Small Explorer program.)

Table 1 Space missions mentioned in this chapter

Mission	Operation period
Stratoscope	several balloon flights 1957 and 1959
OSO 4 (Orbiting Solar Observatory)	1967–69
OSO 6 (Orbiting Solar Observatory)	1969–72
Skylab	1973–74
Spektrostratoskop	balloon flight 1975
OSO 8 (Orbiting Solar Observatory)	1975–78
HRTS (High Resolution Telescope and Spectrograph)	rocket and shuttle flights since 1975
TRC (Transition Region Camera)	rocket flights 1979 and 1980
SMM (Solar Maximum Mission)	1980–89
SOUP (Solar Optical Universal Polarimeter)	experiment on Spacelab 2, 1985
NIXT (Normal-Incidence X-ray Telescope)	rocket flights, e.g. 1993
Yohkoh	since 1991
SOHO (Solar and Heliospheric Observatory)	since 1995
TRACE (Transition Region And Coronal Explorer)	since 1998
Solar B	launch scheduled for 2005

from Skylab (Tousey 1977), provided a new diagnostic, the wavelength dependence of the continuum intensity, since at shorter wavelengths the continuum radiation emanates from higher layers (e.g. Vernazza *et al*. 1973, 1981). The advantage of UV and EUV spectra is that they also contain emission lines belonging to different ions that carry information on the temperature in the solar chromosphere, transition region and corona.

A reliable knowledge of the thermal stratification is fundamental for the accurate determination of elemental abundances. The pioneering work by Russell (1929) and the seminal compilation by Goldberg *et al*. (1960) have been followed by increasingly detailed and accurate determinations of the abundances of ever more elements. The current status of our knowledge of solar abundances (from the solar core to its corona) is discussed in the volume edited by Fröhlich *et al*. (1998), with the photospheric abundances being reviewed therein by Grevesse and Sauval (1998). On the whole these abundances agree surprisingly well with the meteoritic values, although there are some minor deviations and some residual uncertainty. The latter is due partly to the inhomogeneity of the solar atmosphere (discussed in Sections 2.2 and 2.3), which has generally not been taken into account when determining abundances. However, at the level of accuracy currently being achieved such inhomogeneities begin to have a significant effect.

2.2 Convection

It was evident relatively early that a single atmospheric component cannot adequately describe the solar photosphere. The dark sunspots and the bright faculae (bright structures most prominent near the limb), already visible with a small telescope, highlight the need for multiple thermal components. Sunspots and faculae are associated with magnetic activity (Section 2.3), but even the quiet parts of the Sun are known to be inhomogeneous since the discovery by William Herschel of solar granulation, bright structures typically 1000 km in diameter separated by a dark network. Figure 1 shows a snapshot of solar granulation surrounding a sunspot. On a larger scale a bright network (most prominent in radiation coming from chromospheric and transition-region layers) is also known to exist. To account for such regions with different brightness, sets of plane-parallel models have been produced (e.g. Vernazza *et al*. 1981, Fontenla *et al*. 1993). Again, UV spectra taken outside the terrestrial atmosphere have played an important role in constructing such model families.

High-resolution observations and the modelling of spectral lines have shown that at least in the photospheric layers it is mainly inhomogeneities at scales smaller than approximately 1000 km on the Sun that are of physical relevance. For example, faculae, which have sizes of 10^4–10^5 km, are found to be composed of many small magnetic elements, each with a diameter of the order of 100 km.

The major inhomogeneity in photospheric layers is introduced by the granulation, which is the surface signature of overshooting convection. The bright granules identify hot upflowing gas overshooting from the convectively unstable layers below the solar surface into the stably stratified photosphere. These are surrounded by multiply connected cool and hence dark lanes of downflowing gas. Properties of the granulation have been deciphered using data obtained with balloon-borne telescopes (with the Stratoscope, Danielson 1961; and the Spektrostratoskop, Mehltretter 1978), in space (Solar Optical Universal Polarimeter (SOUP), Title *et al*. 1989) and from the ground (Muller 1999).

A particular success have been detailed two- and three-dimensional numerical simulations, that is computations of the radiation hydrodynamics under conditions corresponding as closely as possible to those present on the Sun, based on a minimum of simplifying assumptions. Such simulations have reproduced a wide variety of observations (e.g. Nordlund 1984, Lites *et al*. 1989), so that they are likely to include the main physical ingredients necessary to describe solar granulation. Mainly, however, they have led to a better physical understanding of solar convection and the influence of granulation on, for example, abundance determinations (e.g. Solanki 1998). Both observations and simulations suggest that the vertical velocity associated with granules decreases rapidly with height, while the horizontal velocity becomes increasingly strong, being supersonic over portions of the largest granules. This last fact is one of the rare predictions made by theory in solar physics that have been subsequently confirmed by observations.

An oscillatory velocity component is also present in the photosphere and chromosphere. In the photosphere its power peaks occur at a period of around 5 min, while in the chromosphere the power peak lies near 3 min. The 5 min oscillations are evanescent in the solar atmosphere, but propagate in the solar interior. They are used to probe the subsurface layers of the Sun (helioseismology). The amplitude of the vertical oscillatory velocity increases with increasing height and dominates over the vertical granular flow field at the top of the photosphere.

In addition to granulation three larger scales of convection are known to affect the solar atmosphere, mesogranulation (5–7 Mm in size) discovered by November *et al*. (1981), supergranulation (20–30 Mm) discovered by Simon and Leighton (1964) and giant cells (covering 40° in longitude and less than 10° in latitude) discovered by Beck *et al*. (1998) using Dopplergrams recorded by the Michelson Doppler Interferometer (MDI) on SOHO. Granulation has by far the most readily visible signature in the photosphere, followed by supergranulation, while the influence of the

other scales of convection on the solar atmosphere is so subtle that it can only be detected with the help of special techniques. In addition to revealing giant cells, MDI has also provided the best images of supergranulation at the solar surface (from Doppler shifts). Yet another important contribution of MDI to the study of solar convection has come from the application of local helioseismic techniques to time series of MDI Dopplergrams. Such analyses of the solar oscillation spectrum have provided the first images of supergranular flows below the solar surface (Duvall *et al*. 1997). A comparison between the subsurface supergranulation (reconstructed from MDI local helioseismology) and MDI magnetograms provides direct confirmation of the traditional picture that the magnetic network is located at the convergence points of the (subsurface) supergranules, so that the magnetic features float in the downflow lanes of the supergranules (Duvall and Gizon 2000). This increases the confidence in the results of local helioseismology. The study of solar convection has thus been firmly catapulted into the space age by MDI, after the SOUP paved the way.

Simulations are now starting to move beyond granules to the larger convective cells (Ploner *et al*. 2000). They suggest that the large-scale convective phenomena observed at the surface are driven at or very close to the surface itself and are not due to the ionization of helium in deeper layers, as had earlier been suggested.

2.3 Magnetic fields

The strongest structuring agent of the photosphere besides granulation is the magnetic field. It is concentrated into flux tubes with a field strength of 1–1.5 kG at photospheric levels, but also has a weaker component, which contains the same order of magnitude of flux as the tubes, but only a small fraction of the magnetic energy. This has to do with the fact that whereas the flux is proportional to the field strength the magnetic energy scales with the square of the field strength.

In the photosphere the magnetic flux tubes are nearly vertical and can be considered to be vertical bundles of concentrated magnetic field lines surrounded by nearly field-free gas.

The largest flux tubes, or rather their intersections with the solar surface, are visible as dark sunspots, while the smallest ones are the magnetic elements, groups of which form faculae and the network. Sunspots have a diameter lying in the range 4000–60 000 km. They are dark and are distinguished from the generally smaller pores by the fact that sunspots have two main components, a darker umbra (with an effective temperature $T_{\rm eff} \approx 4500$ K) and a less dark penumbra ($T_{\rm eff} \approx 5500$ K), while pores have basically one, umbra-like component. Sunspots have peak field strengths of 2000–3500 G, increasing with size. The field strength averaged over the whole sunspot (i.e. over the cross-section of the flux tube) is 1000–1500 G and is very close to that found for small-scale magnetic elements, although the latter carry up to 10^6 times less magnetic flux. The field strength averaged over the flux-tube cross-section actually appears to be independent of flux-tube size at all heights in the photosphere (Solanki *et al*. 1999). An explanation for this result has so far not been given.

An additional basic property of sunspots is the presence of an outflow in the penumbra along thin horizontal flux tubes embedded in the generally inclined field. This outflow, termed the Evershed effect (Evershed 1909), decreases with height and finally turns into an inflow at chromospheric heights, where it is called the inverse Evershed effect (St. John 1913; see Solanki 1997 for a review). Observations with the spectrometers UVSP on the Solar Maximum Mission (SMM; Chipman 1981) and more recently with the Solar Ultraviolet Measurement of Emitted Radiation (SUMER) and the Coronal Diagnostics Spectrometer (CDS) on SOHO revealed that in the transition region the inverse Evershed effect continues into the umbra, where it manifests itself as a downflow that can reach supersonic velocities at transition region temperatures, in particular in bright structures called sunspot plumes (Kjeldseth-Moe *et al*. 1988, Brynildsen *et al*. 1999).

At the other end of the size scale of flux tubes are the magnetic elements, whose diameters are close to or less than the best currently achievable spatial resolution (which corresponds to approximately 150 km in visible light). These flux tubes are bright and harbour most of the magnetic energy in photospheric layers, although they probably carry less than half of the total magnetic flux (Meunier *et al*. 1998). They are constantly being moved around by granules and supergranules. Rapid jostling by granules may produce waves propagating upward along the flux tube (Roberts and Ulmschneider 1997, Grossmann-Doerth *et al*. 1998), which may contribute to the heating of the chromosphere and corona (e.g. Choudhuri *et al*. 1993). They also appear to periodically dissolve or fragment and form again later (Berger *et al*. 1998, Gadun *et al*. 2001). Schrijver *et al*. (1998) estimated from the high rate of magnetic flux emergence revealed by MDI magnetograms of the quiet Sun that the flux in the magnetic network is replaced every 40 hours. The flux emerges in the form of a small loop, whose footpoints move ever further apart with time and rarely come together again later. Thus, from the way that the emerged flux is seen to evolve it is clear that reconnection between field lines must be commonplace and must happen almost uninterruptedly (Section 4.2).

Magnetic elements and sunspots not only structure the photosphere, through their field lines they provide also links to the chromosphere and corona. Along these links energy can be transported from the solar interior (where it

is present in abundance) to the outer atmosphere, where it needs to be deposited. Significant advances in the study of photospheric fine structure (granulation, sunspots, magnetic elements) and of its connection to chromospheric and coronal features are expected to follow from the Solar B mission, currently scheduled to be launched in 2005.

Magnetic elements and sunspots are also thought to be largely responsible for the observed fluctuations of the solar irradiance, that is the brightness of the whole solar disk as measured from above the Earth's atmosphere. All successful observations of solar irradiance variations have been carried out from space. In the meantime the irradiance could be monitored almost continuously for two full solar cycles, although with a variety of instruments whose records only partially overlap (Fröhlich 2000). Such measurements have led to the discovery of brightness dips lasting weeks, produced by the passage of sunspots across the solar disk in connection with solar rotation, as well as a brightening (by 0.1% in total irradiance) at solar activity maximum relative to activity minimum (Willson and Hudson 1991). Both these effects can be quantitatively reproduced on the basis of the evolution of the surface area and spatial distribution of sunspots and faculae on the solar surface (Solanki and Fligge 2000).

3 THE CHROMOSPHERE

3.1 The chromospheric spectrum

The solar chromosphere is visible without filters for a short time at the beginning and end of totality of a solar eclipse at which point the solar limb changes colour dramatically. Outside of eclipses it can be observed in the cores of strong absorption lines at visible or near-ultraviolet wavelengths. Alternatively many of the emission lines in the extreme ultraviolet (EUV) part of the spectrum arise in the chromosphere.

Prominent spectral lines of chromospheric origin are the Ca II H and K lines at around 390 nm, or the Mg II h and k lines at around 280 nm. The cores of these absorption lines show a central intensity peak, which indicates a reversal of the temperature gradient with height, that is that the temperature decrease with height in the photosphere is followed by a temperature increase in the chromosphere, although alternative explanations (assuming a time-dependent or spatially structured chromosphere) are also possible (e.g. Carlsson and Stein 1995).

The spectrum of the chromosphere and of the hotter transition region and corona is, however, much richer when observed from space for two reasons. Firstly, the shorter wavelengths at which most of the transitions from ionized species that sample higher temperatures take place can only be observed from above the Earth's atmosphere. Secondly, at increasingly shorter wavelengths the continuum is formed at ever greater heights, so that at wavelenghts below roughly 160 nm all the spectral lines must be formed in the chromosphere or above. Another advantage of the EUV is that in contrast to the cores of strong lines in the visible many of the lines in the EUV are optically thin and thus easier to analyse. It is therefore not surprising that a significant part of the effort in space-based solar physics had been invested into EUV spectroscopy, culminating in the two spectrometers CDS and SUMER on board the SOHO spacecraft. In particular the latter has provided extremely rich spectra of the chromosphere and the transition region at high spatial and spectral resolution. In Figure 3 a spectrum of the quiet Sun obtained by SUMER is plotted (Curdt *et al*. 1999). Note the large number of emission lines in the spectrum. The first- and second-order spectra of the SUMER grating overlap and refer to the lower and upper wavelength scales, respectively. Most prominent are the Lyα line of hydrogen at 1216 Å and the Lyman continuum starting near 912 Å. Figure 4 shows blow-ups of two spectral regions, with the identifications of the main spectral lines being indicated. Many transitions of neutral (C I, N I, Ne I, S I) and singly ionized species (Fe II, N II, C II, Ar II) are visible. Most of these are of chromospheric origin.

3.2 Standard chromospheric models

In standard, time-independent plane-parallel models of the solar atmosphere (such as those of Fontenla *et al*. 1993) the chromosphere covers the height range between the temperature minimum at the top of the photosphere and the bottom of the transition region, where the temperature increases rapidly with height. In the lower chromosphere of such a model the temperature increases outward appreciably but becomes reasonably height independent in the middle and upper chromosphere. The chromosphere is the lowest part of the atmosphere in which the temperature increases significantly outward from the solar surface and which therefore definitely cannot be in radiative equilibrium, but rather requires some mechanical or magnetic source of energy input. The current picture of the solar chromosphere is geometrically far more complex and more dynamic than portrayed by such 'standard' models (Section 3.4), but for many purposes the standard plane-parallel models are adequate and are still used.

The chromospheric layers of these models are in general based on EUV continua below 1600 Å and lines formed partly or completely in the chromosphere (Lyα, Mg II k, Ca II K). Of particular importance have been spectra obtained by the Orbiting Solar Observatories OSO 4 and OSO 6 (Vernazza *et al*. 1973) and by Skylab (Vernazza *et al*. 1981, Fontenla *et al*. 1993). A difficult problem within the context of these models has been posed by the hydrogen Lyα line, whose profile could not be reproduced by static models.

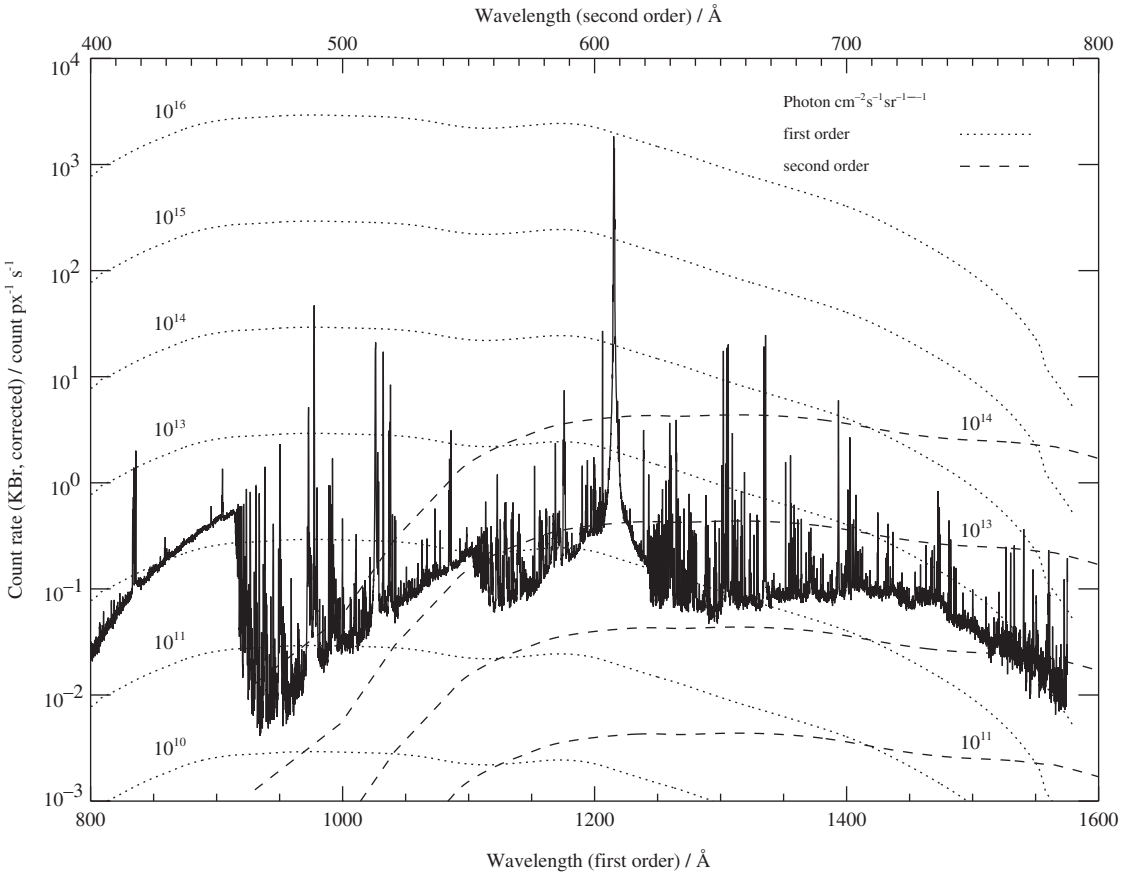

Figure 3 Quiet Sun spectrum of 12 August 1996 from 01:13 to 03:40 UT in first order from 800 Å to 1590 Å. The spectrum is corrected for detector dead time and local gain depression effects, and the attenuation at and near H I Ly α is compensated. Isoradiance contours render the radiometric calibration. (Courtesy of W. Curdt.)

Only the introduction of ambipolar diffusion by Fontenla *et al*. (1993, and reference therein) led to satisfactory results.

The presence of horizontal structuring in the chromosphere, clearly visible in ground-based images made in the core of the Ca II K line and in all chromospheric UV and EUV lines (Bonnet *et al*. 1980, Lemaire *et al*. 1997) is taken into account in a simple manner in these models by introducing a set of plane-parallel atmospheres, each describing different parts of the Sun (different atmospheric components) ranging from the interiors of network cells (darkest and coolest) to the network (brightest and hottest).

3.3 Chromospheric heating

Soon after it was realized that the solar corona is hot, Biermann (1946) suggested an explanation that evolved into the standard heating theory for both the solar chromosphere and corona over a period of three decades. This theory is based on acoustic waves that are generated abundantly in the turbulent flow field of the upper convection zone. As these waves propagate upward, their profile changes since in large-amplitude waves the wave peaks have a higher propagation speed than the valleys. Therefore the peaks attempt to overtake the valleys. This leads to the formation of shocks – thin zones in which the velocity and temperature change so rapidly that viscosity and thermal conduction convert the wave energy into heat, which is then available to sustain the elevated temperature of the upper solar atmosphere.

The formation of shocks out of sound waves is not an everyday experience. Imagine, for example, that we confine the sound of a tuning fork into a tube in order to enforce one-dimensional propagation. Such a small-amplitude wave must travel a distance of the order of the circumference of Earth before it shocks. The waves generated in the upper solar convection zone can form shocks a hundred times faster, despite their longer wavelengths. This is because of their large amplitudes – they are already large when the waves start in the convection zone, and they increase further as the waves propagate upward into a region of decreasing density.

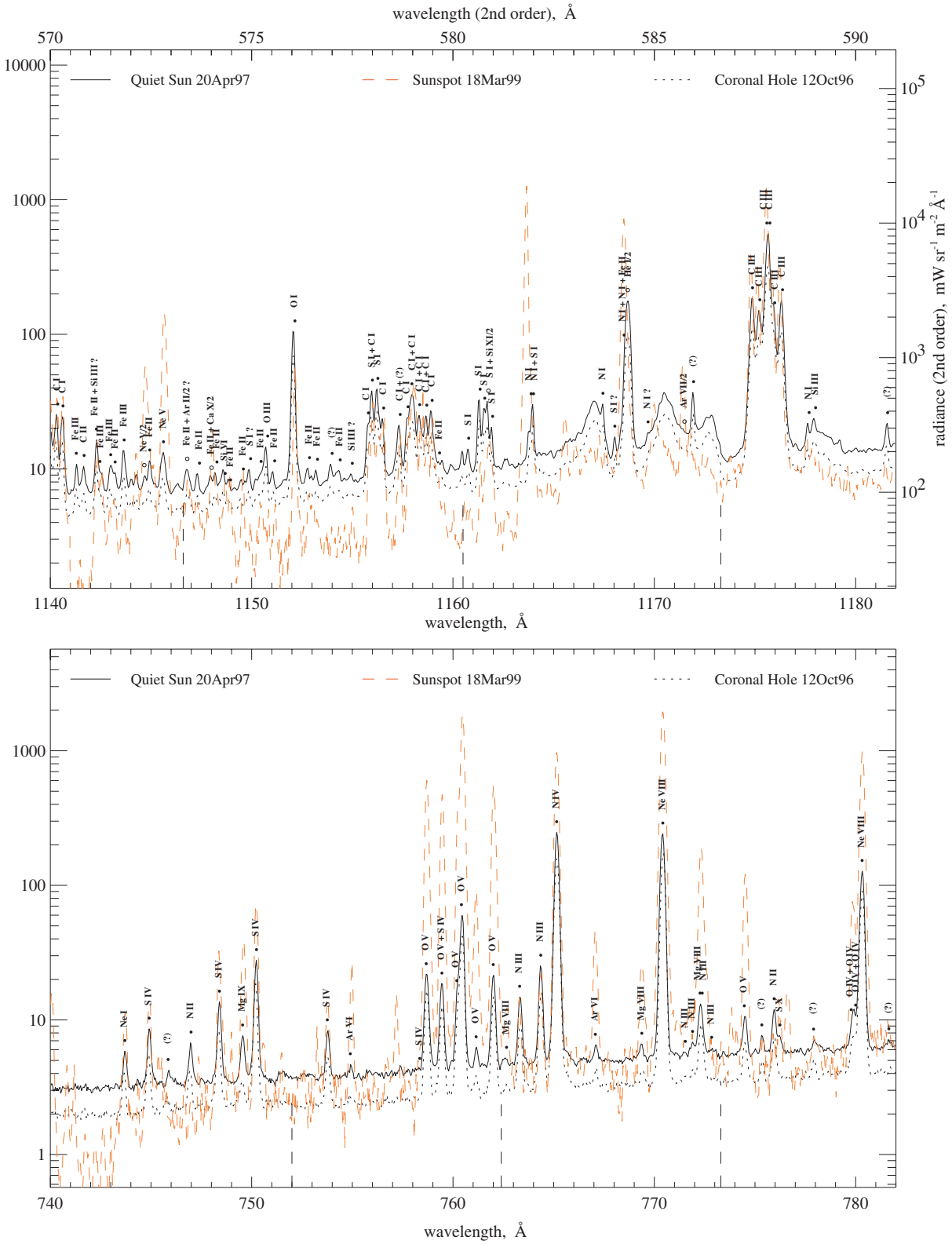

Figure 4 Samples from the spectral atlas in the spectral range 740 to 782 Å and 1140 to 1182 Å, representing quiet Sun (solid), sunspot (dashed), and coronal hole (dotted). (Courtesy of W. Curdt.)

When shocks have been formed, the wave amplitude is controlled by two competing effects: the outward density decrease in the solar atmosphere tends to increase it further, whereas the energy loss associated with shock dissipation tends to decrease it. Ultimately both effects balance each other; then the shock amplitude remains roughly constant. The heating associated with such a wave of constant amplitude is proportional to the density. This matches nicely the behaviour of the radiative output in standard (hot) chromosphere models (Ulmschneider 1970, Anderson and Athay 1989).

Therefore, chromospheric heating by acoustic shock waves has some attractive properties: it starts in the lower chromosphere, after shocks have been formed, and then it decays in a way consistent with the observed radiation loss. For these reasons, acoustic waves are still thought to be the main heating source of nonmagnetic parts of the solar chromosphere. They can also explain the so-called 'basal' emission from the chromospheres of stars with very low magnetic activity (Buchholz *et al*. 1998).

The same properties that make acoustic waves attractive for heating the chromosphere, however, make them unsuited for heating the corona. Their energy flux, decaying proportionally to the density, is virtually exhausted when they reach the corona. This was shown with data obtained by the OSO 8 satellite, which placed very severe constraints on the amount of upward propagating acoustic energy flux in the upper chromosphere and lower transition region (Athay and White 1978, Bruner 1981). Ground-based observations (e.g. Schmieder and Mein 1980) confirmed these constraints, and theoretical models showed that purely acoustic waves are unlikely to be able to produce a corona in either Sun-like or giant stars (Hammer and Ulmschneider 1991).

Sophisticated numerical simulations of shock waves in the nonmagnetic solar chromosphere (Carlsson and Stein 1995, 1997) reproduced very well the observed brightenings in the spectral lines H and K of singly ionized calcium and showed convincingly that the observed time-dependent variations in the line profiles are caused by the passage of large-amplitude shocks. The temperatures and densities behind these shocks were found to be so high that these postshock regions alone could produce the entire emission from the lower chromosphere, without an outward rise of the average chromospheric temperature. However, these simulations neglected short-period acoustic waves, which are also generated in the convection zone and could provide some background heating, as argued by Kalkofen *et al*. (1999). Therefore it is not yet clear if the entire lower chromosphere in nonmagnetic regions is cool and where the average temperature starts to rise. This problem will be discussed further in Section 3.4.

In magnetic parts of the chromosphere, in particular in the network, the magnetic field plays an important role in the heating process. In the simplest case, this role is rather passive, when magnetic flux tubes only act as quasi-static ducts that channel longitudinal wave motions. Such waves can be generated by squeezing flux tubes in convective flows. They can also be generated out of transverse flux tube waves, similar to water splashing out of a water hose that is shaken around. The presence of a magnetic field changes the propagation speed of longitudinal tube waves only slightly; however the onset of shock formation and heating can be delayed in flux tubes that expand rapidly because the wave amplitude grows slower when the wave energy is spread over an increasing area. Otherwise longitudinal tube waves have similar properties as regular acoustic waves; in particular they heat rapidly and are thus good candidates for heating magnetic parts of the chromosphere of the Sun and other stars (Cuntz *et al*. 1998).

The magnetic field can also play a much more active role in the heating process, for example when transverse or torsional wave motions of the magnetic field carry the main wave energy, or when stored magnetic energy is released when different magnetic structures collide. Such heating mechanisms will be discussed further in Section 5.1 in the context of coronal heating. Several of the mechanisms that might heat the corona could also contribute to the heating of magnetic parts of the chromosphere.

3.4 Thermal and dynamic structure of the chromosphere

An image taken in almost any spectral line formed at chromospheric temperatures reveals prominent spatial structure, the most dominant being the so-called chromospheric network. This consists of patches of enhanced brightness located along the boundaries of supergranulation cells (e.g. Bonnet *et al*. 1982). This network is quite inhomogeneous and patchy, with its brightness being related to the amount of magnetic flux concentrated at that particular location. In general, the chromospheric network looks qualitatively similar to that visible in lower transition-region lines, such as He II 304 Å, although the individual network features in the chromosphere are usually finer scale. In the chromosphere the enhanced brightness in the network is thought to be caused by the dissipation of waves travelling along the magnetic flux tubes (cf. Section 3.3). Magnetoacoustic waves, that is acoustic waves modified by the magnetic field and propagating along the field lines are the principal mode thought to be of importance for chromospheric heating. In contrast to other, incompressible, wave modes supported by magnetic flux tubes magnetoacoustic waves with reasonable amplitudes in the photosphere steepen to form shocks at chromospheric heights. Such shocks provide an efficient mechanism for the dissipation of the wave's energy, that is its conversion into thermal energy of the local gas.

Although in atomic lines the brightness contrast between network and cell interior is generally smaller in the chromosphere than in the transition region, there is evidence from molecular transitions that this is not the whole story.

The fundamental band of the rotational–vibrational transitions of the CO molecule is located around 4.8 μm in the IR. When observed near the solar limb or off the limb the cores of the strongest lines of this band are formed in the chromosphere and exhibit temperatures as low as 3500–3800 K, that is below the traditional temperature-minimum value (Ayres and Testermann 1981, Solanki *et al*. 1994). This led to the picture that much of the lower solar chromosphere is in a cool state, quite different from that described by standard chromospheric models (Section 3.2). Combining these observations with the picture revealed by atomic chromospheric lines such as the Ca II H and K lines (which show the bright chromospheric network) it was concluded that only a small fraction of the solar chromosphere is actually in a hot state. This consists mainly of the network, with the cell interiors being very cool according to this picture (Ayres *et al*. 1986). On the other hand, using data obtained by SMM, Athay and Dere (1990) deduced that at least in the layers at which the O I and C I lines at UV wavelengths are formed, 90% of the solar suface is covered by gas at chromospheric temperatures.

The theoretical ideas used to explain this complex and seemingly contradictory thermal structure have evolved considerably with time. Ayres (1981) first pointed out that the CO molecule itself may contribute to the high contrast in temperature between that deduced from atomic lines and the CO lines. If the heating rate is not sufficiently high and the chromospheric temperature drops below a given value (roughly 4000 K), CO begins to form rapidly. Since this molecule is an efficient radiative cooling agent it then lowers the temperature even further until a steady state is reached. Sufficiently above the critical temperature CO never forms and the atmosphere remains hot. This process leads to a thermal bifurcation.

More recently another, more dynamic scenario has emerged from the one-dimensional radiation-hydrodynamic simulations of Carlsson and Stein (1995). These indicate that the passage of shock waves through the chromosphere produces strong peaks in temperature that can be higher than the traditional chromospheric temperature. The chromospheric emission in atomic lines samples mainly these high temperatures, while the CO lines are a better diagnostic of the cool gas between the shocks, since CO gets dissociated at the shock temperatures. Interestingly, when averaged over time the chromospheric gas in these simulations is cool, with temperatures close to those deduced from CO (Section 3.3).

Finally, in three dimensions we expect interactions between acoustic waves propagating in different directions (Rutten and Uitenbroek 1991) as well as overshooting from granular convection in the lower chromosphere (e.g. Steffen and Muchmore 1988) to also play a role in shaping and structuring the chromosphere.

3.5 Magnetic canopy

Magnetograms obtained near the solar limb in chromospheric lines reveal the presence of large patches of almost horizontal magnetic field, which have been interpreted as the base of a magnetic canopy (Giovanelli and Jones 1982). This means that above a certain height, which is thought to be around 700–1000 km in the quiet Sun, the atmosphere is filled with magnetic field. Recently, independent Hanle effect measurements have confirmed the presence of such a magnetic canopy (Bianda *et al*. 1999). The Hanle effect is a quantum interference effect that allows the measurement of much weaker magnetic fields in the solar atmosphere than is possible with the generally used Zeeman effect.

The explanation of such a low-lying magnetic canopy makes use of the large temperature contrast between the hot magnetic elements and their cool surroundings since the pressure scale height is proportional to temperature. The magnetic field in a flux tube is confined by the excess of gas pressure in the surroundings. In the hotter flux tubes the gas pressure decreases less rapidly with height than in the cooler surroundings. Solanki and Steiner (1990) showed that for empirically derived temperature stratifications the internal gas pressure becomes larger than the external value at around 700–1000 km. At this height the magnetic field cannot be confined by the external gas any more and spreads out rapidly. Thus a canopy is formed, in agreement with the observations.

One interesting quantity describing the relative importance of gas and field for the dynamics is the plasma $\beta = 8\pi p/B^2$, which is simply the ratio of the thermal energy density of the gas to the magnetic energy density. In the photospheric layers of flux tubes the magnetic energy density dominates over that of the gas, that is, $\beta < 1$. In the canopy, however, β can locally become larger than unity, since due to the rapid expansion the field becomes rather weak. Higher in the atmosphere (e.g. the corona) $\beta \ll 1$ everywhere and the dynamics are to a large extent magnetically driven. This decrease in β is produced because the gas pressure decreases exponentially, whereas above the canopy the field strength decreases only slowly, roughly following a power law.

4 TRANSITION REGION

The transition region between the chromosphere and corona belongs to the most fascinating parts of the solar atmosphere. It separates two vastly different temperature regimes, in which the energy balance between heating and cooling processes operates in different ways. This thermal interface is thin, but highly structured and extremely dynamic. There

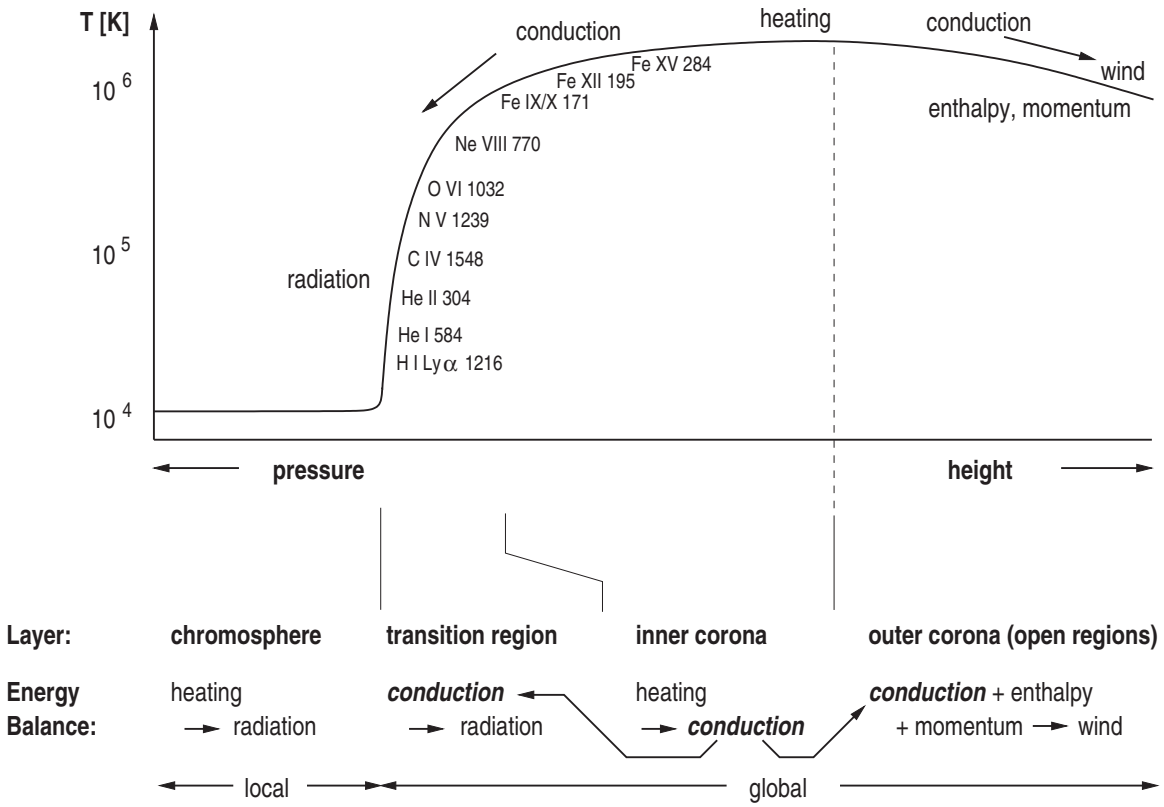

Figure 5 Schematic representation of the energy balance and average temperature along a magnetic flux tube in the outer solar atmosphere. The right-most part of the diagram (labelled outer corona) applies only to 'open' flux tubes that extend out to interstellar space. This part of the diagram does not apply to flux tubes belonging to a coronal magnetic loop; here the remaining part of the diagram describes one loop 'leg', from the solar surface up to the point near the loop top where the maximum temperature is reached. The chromosphere is characterized by a *local* balance between heat input and radiative output. Transition region and corona are characterized by a *global* energy balance, where the heat input is redistributed by thermal conduction to the places where it is needed. In both open and closed flux tubes, heat is conducted back into the inner corona and transition region, where the densities are high enough that the energy can be radiated away in strong emission lines of various ions. In magnetically open regions, heat is also conducted outward and helps to lift the solar wind out of the solar gravitational field and to accelerate it to its high speed. In addition to conduction, the wind is also powered by the energy set free by the cooling of the outflowing gas ('enthalpy') and by the direct energy and momentum input from waves.

are indications that significant variations occur on spatial and temporal scales that are smaller than could be resolved with the best current instruments. In fact, the transition region may be that part of the solar atmosphere that will ultimately impose the highest demands on the spatial and temporal resolution of future space observations if we ever want to understand the dominating physical processes in sufficient detail.

4.1 Energy balance

Before we discuss small-scale variations, a few simple theoretical considerations are in order, to explain the different character of the energy balance in the chromosphere and corona, why these two regimes are separated by a thin transition region and where this transition region is located. The essence of these theoretical considerations is illustrated in Figure 5.

Throughout the entire chromosphere, the Sun is able to adjust the temperature and ionization state of the gas in order to radiate away the energy that is locally deposited by the heating mechanisms. The further out we go, however, the lower the density becomes, and the longer it takes for the atmosphere to cool by radiation after a heating event. Ultimately radiation becomes so inefficient in the outer, tenuous parts of the solar atmosphere that any substantial amount of heating leads to the formation of a hot corona.

In the corona, the temperature is large enough that all atoms are highly ionized, and that the free electrons have such high thermal speeds that they can transport energy very efficiently from hotter to cooler places. This energy redistribution by thermal conduction changes the character of the

energy balance fundamentally, since now the energy sources (where the heating occurs) and sinks (where the radiation is emitted or gas motions are powered) can be spatially separated. It is through this trick of globalizing its energy balance that the corona can once again achieve energy equilibrium. Thermal conduction (and to some extent also flows and waves) collects the heat input in the inner corona and transports it back towards higher densities where it can be radiated away more easily. In magnetically open regions, part of the energy is also transported outward, where it helps to lift the solar wind plasma out of the gravitational field of the Sun and to accelerate it to its final speed.

A major energy sink of the corona is the transition region to the underlying chromosphere. Within this thin layer the temperature jumps by two orders of magnitude, from around 10^4 K in the upper chromosphere to around 10^6 K in the corona (Figure 5). Many ions with strong spectral lines exist in this temperature range. Particularly large amounts of energy are emitted in the resonance lines of the most abundant elements, hydrogen (H I Ly α 1216 Å) and neutral (He I 584 Å) and singly ionized (He II 304 Å) helium, which are all formed at the foot of the transition region. Numerous other strong spectral lines from a variety of ions are also formed in the transition region. Since virtually all of these emission lines are located in the UV and EUV parts of the spectrum, which are absorbed by the Earth's atmosphere, any direct observational information on this layer must come from space instruments.

For the outer solar atmosphere, the total emitted energy per volume and time can be shown to vary as the square of the density times a function of temperature. This function peaks around 10^5 K; and moreover the density in the lower transition region is two orders of magnitude higher than in the corona. This explains why the transition region radiates much more efficiently than the corona. Plasma with a density and temperature typical of the lower transition region can radiate away its thermal energy in a so-called radiative cooling time of only a few seconds, while plasma with properties typical of the inner corona needs about an hour to cool by radiation.

For this reason a lot of energy is transported by thermal conduction from the corona back into the transition region, from where it is radiated away. At high temperatures, thermal conduction is mostly due to electrons, which are bound to follow the magnetic field lines. The conductive energy flux is then given by the temperature gradient along the magnetic field times the thermal conductivity, which is determined by the speed at which the electrons can move. As the conductive energy flows downward to lower temperatures, the electron speed decreases rapidly. The resulting decrease of the conductivity must be offset by a steepening of the temperature gradient in order that the energy can still be transported further down. As a result, the transition region is a thin layer with a particularly steep temperature gradient in its lower parts.

The location of the onset of the transition region within the chromosphere determines its pressure, or density, and thus the amount of energy that it can radiate away. This location is therefore adjusted to changes of the coronal heating rate. Suppose the energy input into the corona is temporally enhanced. Then the excess energy can no longer be radiated away either in the corona or the transition region at their current densities. It is therefore conducted right into the upper chromosphere, where it heats up plasma, which then expands into the corona. This process of 'chromospheric evaporation' effectively pushes the transition region downward until the density in the transition region and corona becomes high enough that the excess heat input can be radiated away. Conversely, when the heat input into the corona is temporally reduced, coronal plasma cools and flows back down into the chromosphere ('coronal condensation'). In this way fluctuations in the coronal heating rate generate flows through the transition region.

4.2 Structure

The transition region is not simply a thin, spherically symmetric shell around the Sun. It is highly structured by the magnetic field. The latter fills all available space already above the middle chromosphere (Section 3.5), but is still distributed inhomogeneously in the overlying transition region and corona. Even along a given magnetic field line the location of the transition region varies in time, in response to the coronal heat input, as discussed in Section 4.1.

On the disk, the most prominent feature seen in spectral lines from the lower transition region (Figure 6) is the network, an extension of the chromospheric network to higher temperatures. The brightest network elements are up to an order of magnitude brighter than the darkest points in the cell interior. However, not the entire network lane area is bright at any given point in time. The total network area accounts for about two-thirds of the total emission. The width of the network lanes is typically around 7 Mm (e.g. Patsourakos *et al*. 1999), while the cell diameters are of the order of 20–30 Mm. Beyond 2.5×10^5 K the network widens rapidly with increasing temperature, hence it becomes increasingly diffuse and is no longer recognizable above 10^6 K (Figure 2).

The ultimate reason for the enhanced emission in the network is that there the magnetic flux density is larger than in the cell interior. Horizontal convective flows transport newly emerging magnetic flux from the interior of supergranulation cells towards the borders, where it accumulates. Studies with the MDI instrument on the SOHO satellite showed that some of these flux elements disappear during this trip, for example by subduction under the solar

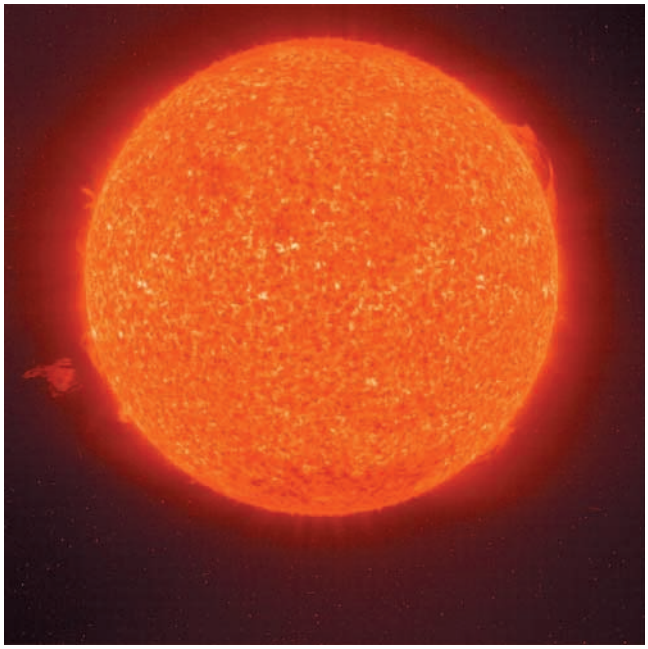

Figure 6 Solar image, taken with the Extreme Ultraviolet Imaging Telescope (EIT, Delaboudinière *et al.* 1995) on SOHO in a wave-length band dominated by the spectral line He II 304 Å of singly ionized helium, formed around 60 000 K in the lower transition region. On the disk, the most characteristic feature seen at these temperatures is the network structure. The limb shows several prominences, which consist of plasma much cooler than the surrounding corona, supported by a magnetic field. These structures can become unstable and erupt, as in the lower left. Other limb features are needle-like EUV macrospicules. The significantly reduced brightness in coronal holes (like the one at the bottom of the image) is not typical of spectral lines formed in the lower transition region, but a special property of helium lines. (Courtesy of the SOHO/EIT team. SOHO is a project of international cooperation between ESA and NASA.)

surface. Others, however, reach the borders of the supergranulation cells, at a rate sufficient to replenish the network magnetic field within 40 hours (Schrijver *et al.* 1998). Some of these magnetic flux elements collide with pre-existing network field, often causing reconnection events (Chapter 43). The latter might be responsible for most of the observed dynamics, which will be discussed later. Reconnection also leads to the reconfiguration of the magnetic field. As a result, the network field is a continuously changing mixture of magnetic flux tubes of various geometries and sizes, consisting mostly of a range of magnetic loops, from the smallest ones that cannot yet be resolved observationally to large ones that connect to the solar surface at large distances. In some areas there are also open flux tubes, which reach out all the way to the interstellar medium. The small loops do not extend to large heights. Thus with increasing height the magnetic structure becomes simpler, and only the so-called coronal funnels (the legs of large loops and open regions) survive and expand into the available space above the cell interior. Models of this expansion (Gabriel 1976) describe the observed widening of the network in the upper transition region quite well. The emission from the lower transition region ($T < 2 \times 10^5$ K), however, is much harder to model. This will be discussed in more detail in Section 4.4.

While with increasing temperature the network fades, other features begin to dominate the appearance of the Sun (Figure 2). They outline the magnetic field structure on larger scales and show us the basic components of the corona: coronal holes are regions where the large-scale magnetic field is unipolar, while active regions consist of complex systems of magnetic loops with enhanced magnetic field strength, temperature and pressure. The corona outside of holes and active regions is called the quiet corona; it is composed of loops of lower temperature and pressure than in active regions, and perhaps some interspersed small open field regions.

Coronal holes are barely noticeable in the emission intensity from temperatures corresponding to the lower transition region. The small-scale magnetic network seems to be similar, although not identical (Huber *et al.* 1974), underneath coronal holes as under the quiet corona. With increasing temperature (beyond 5×10^5 K) the brightness contrast between quiet Sun and coronal holes increases, and the densities and temperature gradients are smaller in coronal holes. Near 10^6 K (Figure 2), finally, no emission is visible any more from coronal holes because their maximum electron temperature is somewhat smaller. Here we see only emission from the diffuse quiet corona (with maximum temperatures in the range $1–2 \times 10^6$ K), active regions (with temperatures of several million kelvin) and compact bright points.

Along the solar limb, several types of inhomogeneities can be seen at chromospheric and transition region temperatures, including prominences and various types of spicules. Prominences are relatively cool clouds of gas (at around chromospheric temperatures) that are embedded in the hot corona and usually supported against gravity by the magnetic field. They are often surrounded by, and interspersed with, plasma emitting at transition region temperatures. Spicules can be seen in strong chromospheric lines as columns of gas protruding out of the solar limb, 10 Mm high and less than 1 Mm thick. They consist of chromospheric matter that is ejected upward at speeds exceeding 20 km s^{-1} and then either falls back or disappears from the visible part of the spectrum after a total lifetime of 5–10 min. Spicules transport much more mass into the corona than is needed by the solar wind, so essentially all of it must flow back. Those spicules that are not seen to fall

back at chromospheric temperatures are probably heated up to a few times 10^5 K before most of their gas flows down again and contributes to the transition region emission. On the other hand, as long as the spicule matter is cool, it absorbs part of the transition region emission at short wavelengths from the solar limb. These emission and absorption contributions by spicules are one reason why the transition region does not appear as an extremely thin spherical shell at the solar limb. Macrospicules are giant versions of spicules within coronal holes; they can be seen in EUV lines formed at temperatures up to 2×10^5 K. Fascinating movies of EUV macrospicule jets have been obtained with Skylab and with EIT and CDS on SOHO. The ultimate cause for the upward ejection of (macro)spicular plasma has not yet been identified, although several possible mechanisms have been proposed (including various types of waves as well as scenarios involving the buffeting of magnetic flux tubes by granular motions in the photosphere).

Unfortunately, the best spatial resolution (down to 1 arcsecond, or 700 km on the Sun) that has so far been achieved with space instruments (like the Transition Region Camera (TRC; Bonnet *et al*. 1980), the High Resolution Telescope and Spectrograph (HRTS; Bartoe and Brueckner 1975), the Normal-Incident X-ray Telescope (NIXT; Golub *et al*. 1990), SOHO/SUMER and TRACE) has turned out to be insufficient to resolve much of the fine structure that we now know to be important for the physics of the transition region. We are still far from understanding the complex magnetic topology of the lower transition region and the structure, acceleration mechanisms and ultimate fate of spicules. While recent space experiments have given us a glimpse of the amazing level of fine structure that governs this part of the solar atmosphere, they also uncovered a number of problems that can be solved only with future observations with even better resolution.

4.3 Dynamics

The transition region is not only complicated because of its fine structure, but also due to various types of motions and a generally high level of temporal variations. This is impressively illustrated by TRACE movies like those distributed on CD-ROM with the review article by Schrijver *et al*. (1999). They show variations and apparent motions on all spatial and temporal scales – including brightening, moving and oscillating magnetic loops; magnetic field reconfigurations with associated mass ejections; or up- and downflowing cool gas in spicules and active-region filaments.

From such movies, however, it is often not easy, or even feasible, to assess the speeds involved – some of the apparent motions might simply be moving wave fronts rather than moving gas. A more direct, but still not always unique, measurement of plasma velocities is based on the analysis of the positions and shapes of spectral lines recorded with instruments such as SO82 B (Bartoe *et al*. 1977) on Skylab, OSO 8, HRTS or the spectrometers SUMER and CDS on SOHO.

Such measurements identify motions both towards and away from the observer. The gas that contributes most to the emission from lower transition region temperatures is predominantly moving towards the Sun (Figure 7, left panel), at speeds reaching a maximum of $10\,\mathrm{km\,s^{-1}}$ for temperatures near 2×10^5 K. The average downflow speed decreases both towards lower and higher temperatures and turns into upward speeds for $T > 5 \times 10^5$ K (Peter and Judge 1999; Figure 7). It has been suggested that the apparent down- and upflows represent the compression regions of downward and upward running acoustic waves that are generated by sudden heating events in magnetic loops. Alternatively the observed flow pattern could be caused by chromospheric gas (like spicules and similar phenomena) that was ejected upward at cooler temperatures and then heated to a few times 10^5 K. In this picture, the main part of the gas falls back towards the Sun and causes the downflow at lower temperatures, while the hottest parts of the gas might further expand into the corona, thus causing the outflow at higher temperatures. Flows initiated by highly asymmetric and temporally variable heating in magnetic loops have also been suggested to contribute to the observed velocities. The ultimate explanation of this phenomenon needs observations with even higher resolution as well as more sophisticated numerical simulations. The upflows at higher temperatures come mostly from network boundaries or intersections of boundaries, in particular in coronal holes, where the large-scale magnetic field is open. Here the measured outflows have been interpreted as a signature of the onset of the fast solar wind (Hassler *et al*. 1999, Peter and Judge 1999, Stucki *et al*. 1999, Wilhelm *et al*. 2000), which has long been known to emanate from coronal holes.

Wave motions have also been identified, either as periodic oscillations in individual spectral lines formed in the network and in active regions, or as a time delay between velocity fluctuations in lines formed in the chromosphere and the transition region. The latter provides evidence for upward travelling waves in the network cell interior (Wikstøl *et al*. 2000), presumably the transition region remnants of strong shocks in the chromosphere (Section 3.3).

The width of transition region spectral lines (with the exception of the strongest ones) is mainly determined by the thermal motion of the atoms and by macroscopic, but spatially unresolved gas motions. The latter, so-called nonthermal motions, are found to reach values up to $30\,\mathrm{km\,s^{-1}}$, with a large scatter, and are even larger in coronal holes than in the quiet Sun. Spectral line shapes are often a superposition of multiple components (Kjeldseth-Moe and Nicolas 1977), most often consisting of a narrow main component and a weaker, but broader component that is slightly shifted in

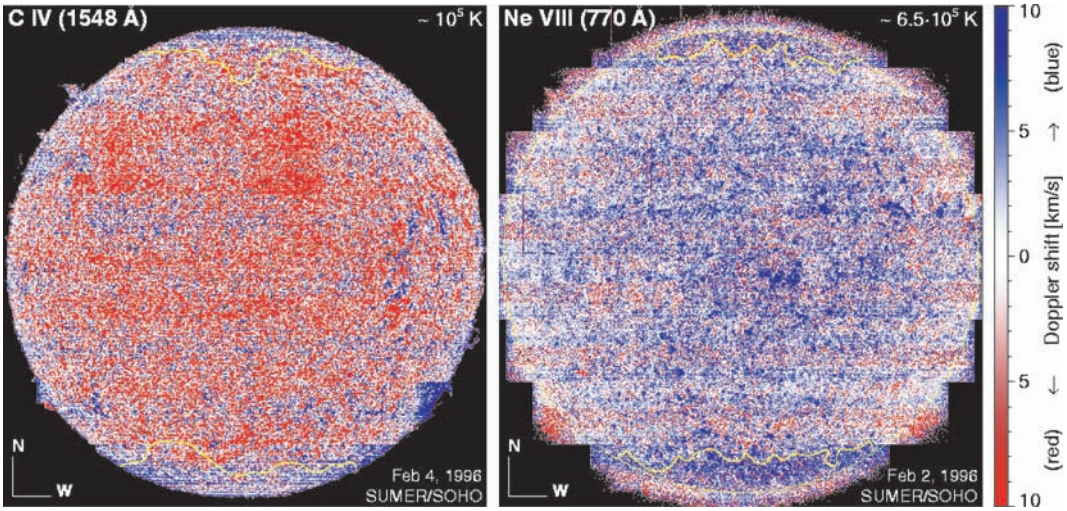

Figure 7 Doppler shifts of spectral lines in the transition region, measured by scanning with the SUMER instrument on SOHO over the Sun. On the disk, emission from lower transition region temperatures (left panel) is dominated by redshifts. They could be produced by flows and downward travelling waves generated in transient, asymmetric heating events in magnetic loops or by down-falling plasma at 10^5 K that has been lifted up at other (presumably lower) temperatures. In the upper transition region (right panel) blueshifts prevail. Coronal holes (bordered by the yellow lines near both poles) are the source regions of the fast solar wind, which is a possible cause of the dominance of blueshifts in these areas. (After Peter 1999.)

wavelength. According to a statistical analysis of SUMER data by Peter (2000), this phenomenon is restricted to bright network elements and can be attributed to the different magnetic structures that exist there: coronal funnels and small-scale magnetic loops (Section 4.2).

Temporal variations in the transition region occur on many different scales. Very common are sudden brightenings in the EUV intensity as detected with CDS and SUMER on SOHO. These events ('blinkers', e.g. Harrison 1997) were identified as density enhancements lasting about half an hour on average. Related brightenings have also been detected with EIT on SOHO. They might represent the reaction of the transition region to energy release events in the corona.

More violent explosive events have been observed with HRTS and SUMER. They occur mostly at network borders, like the brightenings (termed microflares) detected with SMM. Strong lineshifts indicate that gas is ejected from the explosion site as a bidirectional jet with speeds up to $150\,\mathrm{km\,s^{-1}}$ (Innes *et al*. 1997). Simultaneous magnetograph measurements show that explosive events are associated with magnetic field cancellation. Their most likely explanation is therefore that they are generated when small magnetic loops are transported by the convective flow to the network border, where they collide with a pre-existing field in such a way that magnetic field reconnection occurs, which reorganizes the field and converts magnetic energy into heat. At higher temperatures in the corona, Yohkoh (Ogawara *et al*. 1991) has also detected jets associated with reconnection.

These latter types of variability point towards a continuous spectrum of reconnection events, ranging from large flares (a flare is a strong, reconnection-driven explosion; cf. the chapter on solar activity) to microflares and then down to nanoflares, the latter being small-scale heating events along individual field lines within magnetic loops that were suggested to be a basic heating mechanism of coronal loops (Parker 1988). There has been a lot of discussion if the combined effect of all these observed variations is already sufficient to heat the corona in magnetically closed regions. At the time of writing it appears that other heating mechanisms are operating as well (e.g. Aschwanden *et al*. 2000) – but this issue is not yet finally solved, mainly because it is very difficult to estimate the energy released in the smallest events, which are the most common ones.

4.4 Energetics of the lower transition region

As the previous discussion showed, the lower part of the transition region (below 2×10^5 K) plays a special role: it is structured on finer scales and is temporally more variable than the upper transition region, down to scales that lie beyond the capabilities of current instrumentation. Another problem is that the emission from the lower transition region is much more intense than can be explained by simple steady-state transition region models that are based

solely on energy supply by classical electron thermal conduction along the magnetic field (Section 4.1). While such models can well reproduce the average emission of the upper transition region, they fail to explain the observed intensity from plasma below about 2×10^5 K by several orders of magnitude. One reason for this discrepancy is probably that this type of thermal conduction, which dominates in the transition region and inner corona, becomes inefficient at small temperatures, so that the temperature gradient becomes very steep, as explained in Section 4.1. In typical models of this kind, the thickness of the lower transition region is of the order of only 100 km. Such a thin layer does not contain much plasma, and its total emission is thus small – much smaller than observed.

In the real Sun other, more efficient types of conductive energy transport become important in the lower transition region. First, ions and neutral atoms contribute in addition to the electrons. Diffusion lets neutral atoms drift upward to higher temperatures, while ions and electrons drift downward. Second, it is likely that the plasma is strongly filamented by the magnetic field on small spatial scales. If neighbouring filaments have different temperatures, then thermal conduction *across* the magnetic field becomes important at low temperatures (Rabin and Moore 1984). Third, it is possible that the large observed line widths are produced by unresolved turbulent motions, as suggested by some heating theories. The associated energy transport by turbulent eddies enhances drastically the thermal conduction in the lower transition region (Cally 1990). Models incorporating these types of diffusion and enhanced thermal conduction could successfully account for the strong emission at low temperatures.

A number of other possible explanations have been suggested as well. Waves might deposit momentum and energy directly in the lower transition region, thus reducing the need for thermal conduction from above (e.g. Woods *et al*. 1990). Even without direct heating, waves increase the emission due to associated time-dependent effects (Wikstøl *et al*. 1998), although perhaps not to the required amount (Feldman 1998) – but this needs further clarification in terms of more detailed models. Similarly, time-dependent coronal heating also enhances the emission (Athay 1984, Sturrock *et al*. 1990, Roumeliotis 1991).

Moreover, it is important to realize that the fastest electrons have the lowest collision rates with other particles and could thus penetrate the thin transition region with few or no collisions. In particular, fast coronal electrons could reach the lower transition region (Shoub 1983). Even the possibility that the fastest of the chromospheric electrons could, without collisions, produce the entire transition region and corona without any further heating has been discussed (Scudder 1992) and found to produce the correct emission at low temperatures, but to be inconsistent, in its simplest form, with the emission and other properties at higher temperatures (Anderson *et al*. 1996). This latter suggestion needs to be studied further with more detailed theoretical models.

So there exist indeed a number of promising ideas for explaining the observed excess emission at temperatures corresponding to the low transition region (smaller than about 2×10^5 K). The most often discussed suggestion, however, is that a sizable fraction of this emission does not at all originate from a 'transition region' between chromosphere and corona, but rather from an ensemble of magnetic loops that do not reach coronal temperatures and are so small that they could not yet be resolved with current instruments (Feldman 1983, Antiochos and Noci 1986, Dowdy *et al*. 1986; Section 4.2).

The emission from the important spectral lines of helium, the second most abundant element of the Sun, poses additional problems. Both neutral and singly ionized helium requires particularly large amounts of energy to reach electronic states from where lines can be emitted. These large energies can be provided either by collisions with fast electrons in the transition region, or by coronal EUV and X-radiation that ionizes chromospheric helium atoms, which upon subsequent recombination reach the excited energy states from where the lines are emitted.

In relatively weak helium lines we can see through the transition region into the chromosphere, where the latter mechanism dominates. Therefore, such lines are sensitive to radiation from the overlying corona and have thus been used as proxies to infer the amount of coronal X-ray emission, both for the Sun and for other types of stars.

In the strongest helium lines (like the one shown in Figure 6), however, we can usually see only into the transition region, where electron collisions dominate (Andretta and Jones 1997). When compared to other lines excited by electron collisions in the transition region, however, these helium lines are much more intense, by typically an order of magnitude. In fact, the strongest helium line is usually even stronger than the available ionizing EUV and X-ray emission from the corona, which confirms that it is predominantly excited by electron collisions, not by photoionization. Jordan (1975) suggested that the enhancement could be caused by some process that mixes ions with hotter electrons. This would affect the emission from helium more than from other atoms because helium needs more energy for ionization and more time to adjust its ionization equilibrium. Several of the mechanisms discussed above could effectively provide such a mixing of helium ions with hot electrons: diffusion, turbulence, nonthermal electrons or time-dependent effects such as waves or 'bursty' heating. The enhancement factor is observed to increase with increasing emission, from coronal holes to quiet regions, and from network cell interiors to boundaries. This suggests a dependence of the mixing

process on the density, temperature gradient or the magnetic field structure of the overlying corona.

Many details of the formation of the helium spectrum are not yet understood, despite active ongoing research. Recent studies involved instruments like SUMER and CDS on the SOHO satellite, combined with simultaneous ground-based observations or rocket experiments. Several such studies attempted to compare the spatial and temporal behaviour of helium lines with other lines. But the data are complex, and their correct interpretation requires radiative transfer and hydrodynamic calculations, which need further refinements. And the small spatial and temporal scales that characterize the lower transition region call for observations with higher resolution than our current instruments have achieved.

Helium emission is very important for the energetics of the transition region and upper chromosphere, mainly because helium is the most abundant element after hydrogen. Quite surprisingly, its abundance is not constant. Precise values are known only for the solar interior (where the helium to hydrogen abundance ratio is 8.4% by number, as determined from the analysis of solar oscillations) and in the distant solar wind (where the abundances have been measured directly by spacecraft, as discussed in Chapter 47, and where the helium abundance is found to be only about half the surface value, and highly variable). In the entire solar atmosphere, however, the variation of the helium abundance as a function of height and magnetic structure is comparatively uncertain because it is very difficult to determine from spectroscopic observations, in particular in the upper chromosphere and lower transition region, where the important helium lines are formed. It is possible that small-scale abundance variations in space and time contribute to the observed anomalous behaviour of the helium lines.

5 CORONA

There is no unanimously accepted definition of a boundary temperature that separates the upper transition region and corona. An image like Figure 2, which probes plasma at temperatures near 1 million kelvin, illustrates only partial aspects of each of the basic components of the inner corona: coronal holes dominated by an open magnetic field configuration; active-region magnetic loop complexes and their small-scale counterparts, bright points; and finally the quiet corona consisting of more diffuse magnetic loops and small-scale open regions. Measurements from SOHO instruments showed that the electron temperature in coronal holes rises only up to around 8×10^5 K, while from Yohkoh we know that active region loops reach $2–5 \times 10^6$ K during quiet phases, and up to 10^7 K during heating bursts. An image like Figure 2 therefore outlines cooler coronal loops in the quiet corona and in the outer parts of active regions but does not contain significant emission from the coronal parts of coronal holes, while in the hottest parts of active regions it shows only the transition region footpoints of very hot loops. The coronae of magnetically closed and open regions in the solar atmosphere are fundamentally different, not only with respect to their magnetic topology, but also according to their energy balance, temperature, density and other physical properties. Therefore it is useful to discuss them separately.

5.1 Coronal loops

The presence of large magnetic loop structures in the outer solar atmosphere has long been known from ground-based observations such as those of loop prominences at the solar limb, or from images taken in the few coronal spectral lines that happen to fall in the visible part of the spectrum and are thus accessible from the ground. But only after the first rocket-borne X-ray observations, and in particular since the highly successful Skylab mission, did it become evident that essentially the entire X-ray emission from the Sun originates in magnetic loops, and that these loops are not restricted to a few activity centres, but dominate both the active and quiet corona outside of coronal holes. For a review of these early findings see, for example, Vaiana and Rosner (1978). The one-to-one correspondence between magnetic loops and X-ray emission makes X-ray observations an ideal tool for studying magnetic activity on stars other than the Sun (cf. the chapter on stellar interiors and atmospheres).

A series of subsequent space missions led to a continuous refinement of our knowledge of solar coronal loops. Over a complete solar cycle, Yohkoh monitored the properties of loops at high temperatures, to which it is particularly sensitive. The rocket-borne NIXT experiment provided snapshots of the somewhat cooler corona with higher spatial resolution. EIT evolved into a major workhorse on SOHO, taking images of the entire Sun in four different temperature regimes of the transition region and corona, which proved indispensable as synoptic reference for observations with other SOHO instruments. The latter include the spectrographs SUMER and CDS, which measure flow speeds, temperatures and densities in the transition region and inner corona; and the coronagraphs UVCS (Ultraviolet Coronagraph Spectrometer) and LASCO (Large Angle and Spectrometric Coronagraph), which map the structure and flows of the large-scale corona and solar wind. Like EIT, TRACE obtains images at several different temperatures, but with much higher spatial resolution, at which it can observe only smaller sections of the solar disk at a time.

With respect to coronal loops, perhaps the single most important result of all these observations is the appreciation of the high level of structuring and dynamics in the corona.

For example, images with the so far highest resolution show thin magnetic loops that are only one pixel wide along their entire length. It is likely that these fine loops are still underresolved and in reality even thinner. Such images were taken by TRACE in a wavelength band sensitive to radiation from a limited temperature range, usually around 10^6 K. Does the fact that we recognize the complete loop then imply that it has a virtually constant temperature along its entire length? If so, this would impose severe constraints on the distribution of heating along the loop (Schrijver *et al*. 1999) and/or its variation with time. But alternatively it is also possible (Reale and Peres 2000) that even such thin loops consist of a bundle of even thinner magnetic threads, all having their own separate energy balance, temperature profile and time evolution. If this is the case, the emission from the cooler threads would contribute to the TRACE wavelength band predominantly near the centre part of the loop, while the footpoints of the hotter threads would dominate the image near the loop ends. Neither of these different types of threads needs to have constant temperature. Moreover, one should expect that such extremely thin filaments are also temporally variable, so that our observations average in both space and time. Thus the seemingly constant temperature within the observed loops could well be an illusion due to the fact that our spatial and temporal resolution is not yet high enough to resolve the important physical phenomena at their intrinsic scales. Only future improved measurements can help to resolve these issues.

A similar averaging effect could also explain why Yohkoh observes generally thicker loops than TRACE. Yohkoh samples a broader temperature range, to which more threads might contribute (Peres 1999). Another aspect to keep in mind is that Yohkoh senses higher temperatures, where the radiative cooling times are much longer (Section 4.1). Therefore, if the temperatures of individual loop threads are controlled by heating events followed by cooling, one observes at any given instant more threads at high temperatures than at low ones, so one should expect a general trend towards sharper contours at lower temperatures.

In active regions, TRACE observed in its 10^6 K band an irregular pattern of fine-scale structures that resembles, and was therefore termed, 'moss' (see the review by Schrijver *et al*. 1999). This phenomenon is restricted to regions in which simultaneous Yohkoh observations detect hot overlying loops of at least 3×10^6 K. Hence, moss has been interpreted as the upper transition region in hot loop threads. The moss pattern is highly variable, partly because of absorbing cool matter that moves up and down in front of the loop footpoints. This is possible because bright, hot loops have high pressures, which push their transition regions down to height levels below the top of the surrounding chromosphere.

The reason for the overall bright emission from loops is that they keep the gas trapped within them. Since the matter is highly ionized at coronal temperatures, all particles can only spiral around, and move along (but not across), the magnetic field lines. In the inner solar corona, the gas pressure is typically an order of magnitude smaller than the pressure associated with the magnetic field ($\beta \approx 0.1$; Section 3.5), so the gas in a coronal loop cannot normally open up its container (except for the outermost loops, which will be discussed in Section 5.2). As described in Section 4.1, all the energy that is deposited in a loop as heat is ultimately radiated away, after redistribution by thermal conduction and adjustment of the location of the transition region and the loop pressure. The more heat a loop receives, the larger its pressure becomes, and the brighter the emission. By contrast, magnetically open regions can also cool by outward thermal conduction and in particular by energy losses associated with the solar wind, which acts as a 'safety valve'. If a closed and an open flux tube are heated at the same rate, the open one has thus a lower temperature (because of additional cooling mechanisms) and density (because less energy needs to be radiated away) and therefore a much lower pressure.

But what *are* the mechanisms that heat coronal loops? It has become increasingly clear since the late 1970s that acoustic waves cannot transport enough energy beyond the chromosphere (Section 3.3). Therefore, the energy transport into, and very likely also the heating of, the corona must be mediated by the magnetic field. The footpoints of magnetic field lines are anchored in the deep photosphere, where convective and turbulent flows move them around. This gives rise to a large number of different means to transport energy into the upper layers. Unfortunately, we can address here only some of the most basic mechanisms. For comprehensive reviews, see, for example, Narain and Ulmschneider (1990, 1996).

Quite generally, rapid footpoint motions generate waves, which travel upward along the magnetic field; while slow motions let the atmosphere move through a sequence of quasi-equilibrium states, during which electric currents are important.

There exist many different types of waves that travel along the magnetic field. Slender magnetic flux tubes (where the physical quantities can be assumed constant over the cross-section) support three types of waves: *kink* waves when their footpoints are pushed back and forth by the convective motions; *torsional* waves when they are twisted; and longitudinal (or *sausage*) waves when they are squeezed. Additional types of waves are possible in thick magnetic flux tubes, for example modes that run predominantly along the surface of the tube and others that run in the interior. And outside of flux tubes, for example in the space-filling magnetic field above the canopy, there are again different magnetic wave modes possible. All these

types of waves have been investigated in some detail, trying to answer questions like: how much wave energy is generated at which frequency in the convection zone; how does the wave propagate along the structured magnetic field; how much of the wave energy 'leaks' into the surrounding environment and how much is converted into other wave modes; what fraction of the energy is reflected off the transition region; and where and how is the mechanical energy ultimately converted into heat? Those wave types that have compression as their major restoring force, such as longitudinal tube waves, dissipate rapidly, like acoustic waves, and are therefore good candidates for heating the chromospheric network, but not the corona. Only waves where the magnetic field provides the main restoring force are viable candidates for heating the corona.

In addition to these wave heating mechanisms (sometimes called AC mechanisms) there are also several possibilities for non-wave heating caused by slow footpoint motions (DC mechanisms). Newly emerging magnetic flux appears all over the solar surface as small loops, which are subsequently swept by the supergranular flow towards the cell boundaries. When they collide with a pre-existing field in the network, discontinuous changes arise in the directions of neighbouring magnetic field lines. In such a situation the magnetic field can reconfigure itself; and part of the stored magnetic field energy is suddenly set free, causing heating, waves, plasma jets and brightenings. (For a more exhaustive description of such processes see Chapter 43) Reconnection events can also occur when rising freshly emerged flux collides with overlying magnetic flux. Even without collisions between different loops, reconnection might also be important in individual coronal loops. This is because the continuous shuffling of magnetic field line footpoints causes a 'braiding' of the threads that compose the loop, and thus again differences in the direction of neighbouring magnetic field lines, associated with electric currents that may contribute to coronal heating (Parker 1979). If the braiding gets strong enough, the magnetic field again reconfigures itself in reconnection events called nanoflares by Parker (1988). TRACE has indeed observed indications of such a braiding of individual loop threads (Schrijver *et al*. 1999).

Braiding can lead to a state of small-scale turbulence in the corona, which can be maintained by photospheric motions (e.g. Heyvaerts and Priest 1992). In such a state, the conversion of mechanical energy into heat is much more efficient. New results from SOHO and TRACE indicate that at least parts of the corona are indeed in a state of turbulence. For example, TRACE observed coronal loops that oscillated back and forth a couple of times after being hit by a strong wave from a nearby flare. From the damping rate of this oscillation, Nakariakov *et al*. (1999) concluded that the corona is turbulent.

5.2 The open corona

Higher up in the atmosphere the magnetic structure becomes less complex, and magnetic loops disappear. The solar wind opens up and carries away the outermost loop field lines, leading to a configuration with a thin layer with oppositely directed magnetic field on both sides, which contains electric currents. The resulting helmet- and ray-shaped structures are called streamers (Figures 8 and 9). At solar activity minimum, they are mainly confined to the equatorial zone, while with increasing solar activity the streamer belt extends to higher latitudes and becomes more irregular. Except during high-activity phases, large coronal holes are usually located at the polar regions. Their field lines bend around the streamers and occupy at larger distances the entire heliosphere outside of a narrow zone near the equatorial plane (cf. Chapter 47).

The solar wind has been predicted theoretically and then measured by numerous satellite missions (Chapters 9 and 47). One of the main objectives of the SOHO mission was to identify the coronal origins of the solar wind and to study its initial acceleration. Coronal holes have long been identified as the main source of the fast, regular solar wind that dominates the heliosphere at higher latitudes. Coronal holes are structured by plumes, ray-shaped features that can be traced out to several solar radii. They have higher densities and lower temperatures than the interplume regions. Model calculations as well as observations with SOHO instruments (e.g. Wilhelm *et al*. 1998) have now firmly established that the outflow speeds are much lower within plumes, so that interplume regions are the genuine source regions of the fast solar wind. Therefore, the plasma that makes up the fast wind flows first through open funnels within the network (Section 4.2) and is then accelerated within the low-density interplume lanes.

The slower, more irregular wind near the equatorial plane could originate in several different types of solar regions (e.g. Wang 1994, Noci *et al*. 1997), the relative importance of which is still being debated and might depend on the phase in the solar cycle. The possible source regions of the slow wind include (i) the borders of coronal holes, with their larger areal expansion out to large distances (Figure 9); (ii) small open regions between streamers, with their early rapid expansion and subsequent constriction; and (iii) the outermost parts of closed regions, which occasionally open up. SOHO/UVCS measurements indicate that the legs of streamers have similar ion abundances as were measured by satellites in the slow-wind zone.

The SOHO coronagraphs discovered that the slow solar wind is accelerated slowly, while the fast wind is accelerated surprisingly rapidly (e.g. Antonucci 1999, and references therein). Time sequences from the LASCO coronagraph show elongated blobs of gas that appear above

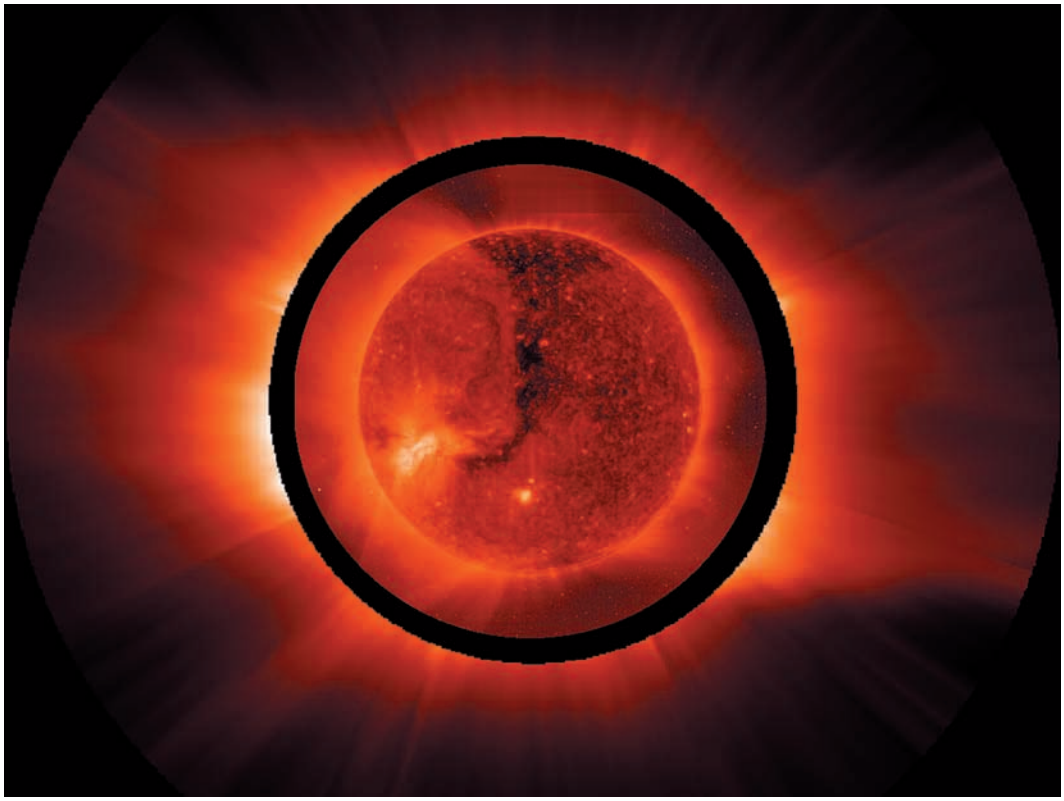

Figure 8 Composite of two SOHO images of the Sun and its corona. The inner part was taken with the EIT instrument and shows the solar disk and inner corona in a wavelength band dominated by iron lines formed near 2 million kelvin. Notice the long extension of the polar coronal hole, two active regions and the streamer regions of enhanced emission over parts of the solar limb. The streamers extend far away from the Sun, as shown in the outer image, taken with the UVCS instrument in the light of fivefold ionized oxygen (O VI). The overall structure of the corona is obviously controlled by the magnetic field. (Courtesy of the SOHO/EIT and SOHO/UVCS teams. SOHO is a project of international cooperation between ESA and NASA.)

the cusps of streamers and are then carried away by the slow wind. They can serve as tracers for measuring the wind speed. One finds that the slow wind is continuously accelerated throughout at least the first 20 solar radii ($R_\odot$), reaching speeds around $300\,\mathrm{km\,s^{-1}}$ there, about three-quarters of the average speed measured at the orbit of Earth. The observed acceleration is consistent with thermal wind expansion (Parker 1958) at a temperature of $1.1 \times 10^6\,\mathrm{K}$. By contrast, observations covering the radial range 1.6–$3\,R_\odot$ in coronal holes indicate that the fast solar wind is accelerated an order of magnitude more rapidly; presumably it reaches its final speed of up to $800\,\mathrm{km\,s^{-1}}$ already within the first $10\,R_\odot$.

It has long been known that the decrease of the density with increasing distance from the Sun must ultimately lead to a situation where collisions between different particles are so inefficient that electrons, protons and other species each have their own temperature and their own energy and momentum balance.

A major result of the SOHO mission was the discovery that the electron temperatures and densities in coronal holes are significantly lower than expected; hence the decoupling between species occurs much earlier than anticipated. The electrons were found to reach their maximum temperature of somewhat less than $10^6\,\mathrm{K}$ already low in the corona, some 10–20% of $R_\odot$ above the photosphere. Beyond this point, the electron temperature declines outward, more rapidly inside plumes than outside. As a result of this low maximum and early decline, and of the inefficient collisional coupling, outward electron heat conduction is lower than expected and may not be able to provide significant amounts of energy out to distances where the fast solar wind needs it: of the order of $R_\odot$ for its potential energy and (as just discussed) a few $R_\odot$ for its kinetic energy. Therefore, much of the wind energy must be supplied either by direct energy input from waves or by enthalpy redistribution (Section 4.1) – that is, by using up the energy from cooling particles that were heated to high temperatures at lower distances.

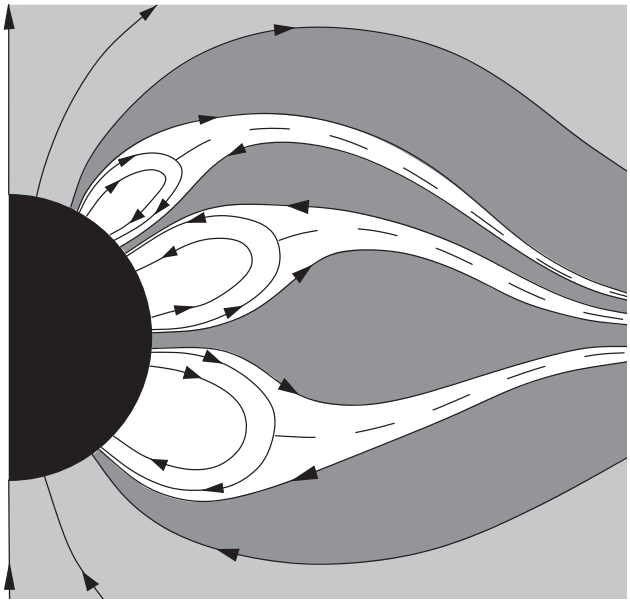

Figure 9 Schematic representation of the large-scale magnetic field structure of the solar corona. The direction of the magnetic field is indicated by arrows. Closed regions exist only up to a certain radial distance, beyond which their field lines are opened up by the solar wind. This leads to regions with oppositely directed magnetic field lines close to each other, and thus to current sheets (dashed lines) within streamers (white). The fast steady solar wind (light grey) comes mainly from large coronal holes, whereas the irregular slow solar wind has been suggested to come from small open regions with strong areal expansion and subsequent constriction (dark grey), with possible further contributions from the boundaries of coronal holes and from the opening of the outermost closed regions.

Both these possibilities appear to play a role. The temperature of protons and neutral hydrogen atoms increases to around $3–4 \times 10^6$ K near $3\,R_\odot$, where both species decouple and hydrogen starts to cool, while the proton temperature can no longer be determined (see the review by Antonucci 1999). Five-fold ionized oxygen ions were also measured with UVCS on SOHO. Their temperature decouples from that of the protons already at $1.5\,R_\odot$ and then increases to extremely high values. Moreover, the velocity distribution of these ions along the magnetic field lines differs markedly from the velocity associated with their gyration motions around the field lines. When interpreted as kinetic temperature, these velocity distributions correspond to at least 10^7 K along the magnetic field direction, and 10^8 K perpendicular to it.

The inefficiency of collisions, and the fact that the temperatures of minor ions are much higher and anisotropic, has been attributed to ion-cyclotron heating by high-frequency Alfvén waves. Such waves can pump magnetic wave energy into the gyration motions of ions. The required waves have been suggested to be generated during reconnection events in the network (Axford and McKenzie 1997), but it is still debated if they can reach far enough into the corona or if they are damped so rapidly that they must be replenished continuously within the corona from the turbulent decay of other waves (Isenberg and Hollweg 1983, Cranmer *et al*. 1999).

REFERENCES

Anderson, L.S. and Athay, R.G. (1989). Chromospheric and coronal heating. *Astrophysical Journal*, **336**, 1089–1091.

Anderson, S.W., Raymond, J.C. and van Ballegooijen, A.A. (1996). Ultraviolet emission-line intensities and coronal heating by velocity filtration: Collisionless results. *Astrophysical Journal*, **457**, 939–948.

Andretta, V. and Jones, H.P. (1997). On the role of the solar corona and transition region in the excitation of the spectrum of neutral helium. *Astrophysical Journal*, **489**, 375–394.

Antiochos, S.K. and Noci, G. (1986). The structure of the static corona and transition region. *Astrophysical Journal*, **301**, 440–447.

Antonucci, E. (1999). Solar wind acceleration region. In J.-C. Vial and B. Kaldeich-Schürmann (eds), *Plasma Dynamics and Diagnostics in the Solar Transition Region and Corona* (Proceedings of the 8th SOHO Wksp.), ESA-SP 446, pp. 53–60.

Aschwanden, M.J., Tarbell, T.D., Nightingale, R.W., Schrijver, C.J., Title, A.M., Kankelborg, C.C., Martens, P. and Warren, H.P. (2000). Time variability of the "quiet" Sun observed with TRACE. II. Physical parameters, temperature evolution, and energetics of EUV nanoflares. *Astrophysical Journal*, **535**, 1047–1065.

Athay, R.G. (1984). The origin of spicules and heating of the lower transition region. *Astrophysical Journal*, **287**, 412–417.

Athay, R.G. and Dere, K.P. (1990). Bifurcation in the low chromosphere. *Astrophysical Journal*, **358**, 710–717.

Athay, R.G. and White, O.R. (1978). Chromospheric and coronal heating by sound waves. *Astrophysical Journal*, **226**, 1135–1139.

Axford, W.I. and McKenzie, J.F. (1997). The solar wind. In J.R. Jokipii, C.P. Sonett, M.S. Giampapa (eds), *Cosmic Winds and the Heliosphere*, University of Arizona Press, Tucson, pp. 31–66.

Ayres, T.R. (1981). Thermal bifurcation in the solar outer atmosphere. *Astrophysical Journal*, **244**, 1064–1071.

Ayres, T.R. and Testerman, L. (1981). Fourier transform spectrometer observations of solar carbon monoxide. I. The fundamental and first overtone bands in the quiet Sun. *Astrophysical Journal*, **245**, 1124–1140.

Ayres, T.R., Testerman, L. and Brault, J.W. (1986). Fourier transform spectrometer observations of solar carbon monoxide. II. Simultaneous cospatial measurements of the fundamental and first-overtone bands, and Ca II K, in quiet and active regions. *Astrophysical Journal*, **304**, 542–559.

Bartoe, J.-D.F. and Brueckner, G.E. (1975). New stigmatic, coma-free, concave-grating spectrograph. *Journal of the Optical Society of America*, **65**, 13.

Bartoe, J.-D.F., Brueckner, G.E., Purcell, J.D. and Tousey, R. (1977). Extreme ultraviolet spectrograph ATM experiment S082B. *Applied Optics*, **16**, 879–886.

Beck, J.G., Duvall, T.S. and Scherrer, P.H. (1998). Long-lived giant cells detected at the surface of the Sun. *Nature*, **394**, 653–655.

Berger, T.E., Löfdahl, M.G., Shine, R.A. and Title, A.M. (1998). Measurements of solar magnetic element dispersal. *Astrophysical Journal*, **506**, 439–449.

Bianda, M., Stenflo, J.O. and Solanki, S.K. (1999). Hanle effect observations with the Ca I 4227 Å line. *Astronomy and Astrophysics*, **350**, 1060–1070.

Biermann, L. (1946). Zur Deutung der chromosphärischen Turbulenz und des Exzesses der UV-Strahlung der Sonne. *Naturwissenschaften*, **33**, 118–119.

Bonnet, R.M., Bruner, E.C. Jr, Acton, L.W., Brown, W.A. and Decaudin, M. (1980). High-resolution Lyman-alpha filtergrams of the Sun. *Astrophysical Journal*, **237**, L47–L50.

Bonnet, R.M., Bruner E.C., Jr., Acton, L.W., Brown, W.A., Decaudin, M. and Foing, B. (1982). Rocket photographs of fine structure and wave patterns in the solar temperature minimum. *Astronomy and Astrophysics*, **111**, 125–129.

Bruner, E.C., Jr (1981). OSO 8 observational limits to the acoustic coronal heating mechanism. *Astrophysical Journal*, **247**, 317–324.

Brynildsen, N., Maltby, P., Brekke, P., Haugan, S.V.H. and Kjeldseth-Moe, O. (1999). SOHO observations of the structure and dynamics of sunspot region atmospheres. *Solar Physics*, **186**, 141–191.

Buchholz, B., Ulmschneider, P. and Cuntz, M. (1998). Basal heating in main-sequence stars and giants: Results from monochromatic acoustic wave models. *Astrophysical Journal*, **494**, 700–714.

Cally, P.S. (1990). Turbulent thermal conduction in the solar transition region. *Astrophysical Journal*, **355**, 693–699.

Carlsson, M. and Stein, R.F. (1995). Does a non-magnetic solar chromosphere exist? *Astrophysical Journal*, **440**, L29–L32.

Carlsson, M. and Stein, R.F. (1997). Formation of solar calcium H and K bright grains. *Astrophysical Journal*, **481**, 500–514.

Chalonge, D. and Kourganoff, V. (1946). Recherches sur les spectres continus du soleil. *Annals of Astrophysics*, **9**, 69.

Chipman, E.G. (1981). The Solar Maximum Mission. *Astrophysical Journal*, **244**, L113–L115.

Choudhuri, A.R., Dikpati, M. and Banerjee, D. (1993). Energy transport to the solar corona by magnetic kink waves. *Astrophysical Journal*, **413**, 811–825.

Cranmer, S.R., Field, G.B. and Kohl, J.L. (1999). Spectroscopic constraints on models of ion cyclotron resonance heating in the polar solar corona and high-speed solar wind. *Astrophysical Journal*, **518**, 937–947.

Cuntz, M., Ulmschneider, P. and Musielak, Z.E. (1998). Self-consistent and time-dependent magnetohydrodynamic chromosphere models for magnetically active stars. *Astrophysical Journal*, **493**, L117–L129.

Curdt, W., Brekke, P., Schühle, U., Wilhelm, K. and Dwivedi, B.N. (1999). The SUMER EUV atlas in the spectral range 670 Å–1498 Å. In: *Proceedings of the 8th SOHO Workshop Plasma Dynamics and Diagnostics in the Solar Transition Region and Corona*, Paris. ESA SP–446. Noordwijk: ESA Publ. Div., pp. 251–256.

Danielson, R.E. (1961). The structure of sunspot penumbras. I. Observations. *Astrophysical Journal*, **134**, 275–288.

Delaboudinière, J.-P., Artzner, G.E., Brunaud, J., Gabriel, A.H., Hochedez, J.F., Millier, F., Song, X.Y., Au, B., Dere, K.P., Howard, R.A., Kreplin, R., Michels, D.J., Moses, J.D., Defise, J.M., Jamar, C., Rochus, P., Chauvineau, J.P., Marioge, J.P., Catura, R.C., Lemen, J.R., Shing, L., Stern, R.A., Gurman, J.B., Neupert, W.M., Maucherat, A., Clette, F., Cugnon, P. and van Dessel, E.L. (1995). EIT: Extreme-Ultraviolet Imaging Telescope for the SOHO Mission. *Solar Physics*, **162**, 291–312.

Dowdy, J.F., Jr, Rabin, D. and Moore, R.L. (1986). On the magnetic structure of the quiet transition region. *Solar Physics*, **105**, 35–45.

Duvall, T.L. Jr and Gizon, L. (2000). Time–distance helioseismology with *f* modes as a method for measurement of near-surface flows. *Solar Physics*, **192**, 177–191.

Duvall, T.L. Jr, Kosovichev, A.G., Sherrer, P.H., Bogart, R.S., Bush, R.I., De Forest, C., Hoeksema, J.T., Schou, J., Saba, J.L.R., Tarbell, T.D., Title, A.M., Wolfson, C.J. and Milford, P.N. (1997). Time–distance helioseismology with the MDI instrument: Initial results. *Solar Physics*, **170**, 63–73.

Evershed, J. (1909). Radial movement in sun-spots. *Monthly Notices of the Royal Astronomical Society*, **69**, 454–457.

Feldman, U. (1983). On the unresolved fine structures of the solar atmosphere in the 30,000–200,000 K temperature region. *Astrophysical Journal*, **275**, 367–373.

Feldman, U. (1998). On the unresolved fine structures of the solar atmosphere. III. Elemental abundances consideration. *Astrophysical Journal*, **507**, 974–977.

Fleck, B. and Domingo, V. (eds) (1995). *The SOHO Mission*, Kluwer, Dordrecht.

Fontenla, J.M., Avrett, E.H. and Loeser, R. (1993). Energy balance in the solar transition region. III. Helium emission in hydrostatic, constant-abundance models with diffusion. *Astrophysical Journal*, **406**, 319–345.

Fröhlich, C. (2000). Observations of irradiance variations. In E. Friis-Christensen, C. Fröhlich, J.D. Haigh, M. Schüssler, R. von Steiger (eds), *Solar Variability and Climate*, Kluwer, Dordrecht. [*Space Science Reviews*, **94**, 15–24.]

Fröhlich, C., Huber, M.C.E., Solanki, S.K. and von Steiger, R. (eds) (1998). *Solar Composition and its Evolution – From Core to Corona*. Space Science Series of ISSI, Kluwer, Dordrecht. [*Space Science Reviews*, 85.]

Gabriel, A.H. (1976). A magnetic model of the solar transition region. *Philosophical Transactions of the Royal Society A*, **281**, 339–352.

Gadun, A.S., Solanki, S.K., Sheminova, V.A. and Ploner, S.R.O. (2001). A formation mechanism of magnetic elements in regions of mixed polarity. *Solar Physics*, **203**, 1–7.

Giovanelli, R.G. and Jones, H.P. (1982). The three-dimensional structure of atmospheric magnetic fields in two active regions. *Solar Physics*, **79**, 267–278.

Goldberg, L., Müller, E.A. and Aller, L.H. (1960). The abundances of the elements in the solar atmosphere. *Astrophysical Journal Supplement Series*, **5**, 1–137.

Golub, L., Nystrom, G., Herant, M., Kalata, K. and Lovas, I. (1990). Sub-arcsecond observations of the solar X-ray corona. *Nature*, **344**, 842–844.

Grevesse, N. and Sauval, A.J. (1998). Standard solar composition. *Space Science Reviews*, **85**, 161–174.

Grossmann-Doerth, U., Schüssler, M. and Steiner, O. (1998). Convective intensification of solar surface magnetic fields: Results of numerical experiments. *Astronomy and Astrophysics*, **337**, 928–939.

Hammer, R. and Ulmschneider, P. (1991). On the intrinsic difficulty of producing stellar coronae with acoustic waves. In P. Ulmschneider, E.R. Priest and R. Rosner (eds), *Mechanisms of Chromospheric and Coronal Heating*, Springer-Verlag, Berlin, pp. 344–346.

Handy, B.N. and 47 coauthors (1999). The transition region and coronal explorer. *Solar Physics*, **187**, 229–260.

Harrison, R.A. (1997). EUV blinkers: The significance of variations in the extreme ultraviolet quiet sun. *Solar Physics*, **175**, 467–485.

Hassler, D.M., Dammasch, I.E., Lemaire, P., Brekke, P., Curdt, W., Mason, H.E., Vial, J.-C. and Wilhelm, K. (1999). Solar wind outflow and the chromospheric magnetic network. *Science*, **283**, 810.

Heyvaerts, J. and Priest, E.R. (1992). A self-consistent turbulent model for solar coronal heating. *Astrophysical Journal*, **390**, 297–308.

Holweger, H. (1967). Ein empirisches Modell der Sonnenatmosphäre mit lokalem thermodynamischen Gleichgewicht. *Zeitschrift für Astrophysik*, **65**, 365–417.

Huber, M.C.E., Foukal, P.V., Noyes, R.W., Reeves, E.M., Schmahl, E.J., Timothy, J.G., Vernazza, J.E. and Withbroe, G.L. (1974). Extreme-ultraviolet observations of coronal holes – Initial results from Skylab. *Astrophysical Journal*, **194**, L115–L118.

Innes, D.E., Inhester, B., Axford, W.I. and Wilhelm, K. (1997). Bi-directional plasma jets produced by magnetic reconnection on the Sun. *Nature*, **386**, 811–813.

Isenberg, P.A. and Hollweg, J.V. (1983). On the preferential acceleration and heating of solar wind heavy ions. *Journal of Geophysical Research*, **88**, 3923–3935.

Jordan, C. (1975). The intensities of helium lines in the solar EUV spectrum. *Monthly Notices of the Royal Astronomical Society*, **170**, 429–440.

Kalkofen, W., Ulmschneider, P. and Avrett, E.H. (1999). Does the Sun have a full-time chromosphere? *Astrophysical Journal*, **521**, L141–L144.

Kjeldseth-Moe, O., Brynildsen, N., Brekke, P., Engvold, O., Maltby, P., Bartoe, J.-D.F., Brueckner, G.E., Cook, J.W., Dere, K.P. and Soker, D.G. (1988). Gas flows in the transition region above sunspots. *Astrophysical Journal*, **334**, 1066–1075.

Kjeldseth-Moe, O. and Nicolas, K.R. (1977). Emission measures, electron densities, and nonthermal velocities from optically thin UV lines near a quiet solar limb. *Astrophysical Journal*, **211**, 579–586.

Lemaire, P., Wilhelm, K., Curdt, W., Schühle, U., Marsch, E., Poland, A.I., Jordan, S.D., Thomas, R.J., Hassler, D.M., Vial, J.C., Kühne, M., Huber, M.C.E., Siegmund, O.H.W., Gabriel, A., Timothy, J.G. and Grewing, M. (1997). First results of the SUMER telescope and spectrometer on SOHO. *Solar Physics*, **170**, 105–122.

Lites, B.W., Nordlund, Å. and Scharmer, G.B. (1989). Constraints imposed by very high resolution spectra and images on theoretical simulations of granular convection. In R.J. Rutten and G. Severino (eds), *Solar and Stellar Granulation*, Reidel, Dordrecht, pp. 349–357.

Mehltretter, J.P. (1978). Balloon-borne imagery of solar granulation. II. The lifetime of solar granulation. *Astronomy and Astrophysics*, **62**, 311–316.

Meunier, N., Solanki, S.K. and Livingston, W.C. (1998). Infrared lines as probes of solar magnetic features. XIII. The relative flux in weak and strong quiet-Sun magnetic fields. *Astronomy and Astrophysics*, **331**, 771–781.

Muller, R. (1999). The solar granulation. In A. Hanslmeier and M. Messerotti (eds), *Motions in the Solar Atmosphere*, Kluwer, Dordrecht, pp. 35–70.

Nakariakov, V.M., Ofman, L., DeLuca, E.E., Roberts, B. and Davila, J.M. (1999). TRACE observations of damped coronal loop oscillations: Implications for coronal heating. *Science*, **285**, 862–864.

Narain, U. and Ulmschneider, P. (1990). Chromospheric and coronal heating mechanisms. *Space Science Reviews*, **54**, 377–445.

Narain, U. and Ulmschneider, P. (1996). Chromospheric and coronal heating mechanisms. II. *Space Science Reviews*, **75**, 453–509.

Noci, G. and 24 coauthors (1997). The quiescent corona and slow solar wind. In A. Wilson (ed), *The Corona and Solar Wind Near Minimum Activity (Proceedings of the 5th SOHO Wksp.)*, ESA SP-404. Noordwijk: ESA Publ. Div. pp. 75–84.

Nordlund, Å. (1984). Modelling of small-scale dynamical processes: Convection and wave generation. In S.L. Keil (ed), *Small-Scale Dynamical Processes in Quiet Stellar Atmospheres*, National Solar Observatory, Sunspot, NM, pp. 181–221.

November, L.J., Toomre, J., Gebbie, K.B. and Simon, G.W. (1981). The detection of mesogranulation on the Sun. *Astrophysical Journal*, **245**, L123–L126.

Ogawara, Y., Takano, T., Kato, T., Kosugi, T., Tsuneta, S., Watanabe, T., Kondo, I. and Uchida, Y. (1991). The Solar-A mission – An overview. *Solar Physics*, **136**, 1.

Parker, E.N. (1958). Dynamics of the interplanetary gas and magnetic fields. *Astrophysical Journal*, **128**, 664–676.

Parker, E.N. (1979). *Cosmical Magnetic Fields*, Clarendon Press, Oxford, pp. 359–391.

Parker, E.N. (1988). Nanoflares and the solar X-ray corona. *Astrophysical Journal*, **330**, 474–479.

Patsourakos, S., Vial, J.-C., Gabriel, A.H. and Bellamine, N. (1999). Transition-region network boundaries in the quiet Sun: Width variation with temperature as observed with CDS on SOHO. *Astrophysical Journal*, **522**, 540–546.

Peres, G. (1999). Coronal heating and structuring. In J.-C. Vial and B. Kaldeich-Schürmann (eds), *Plasma Dynamics and Diagnostics in the Solar Transition Region and Corona (Proceedings of the 8th SOHO Wksp.)*, ESA SP-46. Noordwijk: ESA Publ. Div. pp. 43–52.

Peter, H. (1999). Analysis of transition-region emission-line profiles from full-disk scans of the Sun using the SUMER instrument on SOHO. *Astrophysical Journal*, **516**, 490–504.

Peter, H. (2000). Multi-component structure of solar and stellar transition regions. *Astronomy and Astrophysics*, **360**, 761–776, and erratum in *Astronomy and Astrophysics*, **364**, 933–934.

Peter, H. and Judge, P.G. (1999). On the Doppler shifts of solar ultraviolet emission lines. *Astrophysical Journal*, **522**, 1148–1166.

Ploner, S.R.O., Solanki, S.K. and Gadun, A.S. (2000). Is solar mesogranulation a surface phenomenon? *Astronomy and Astrophysics*, **356**, 1050–1054.

Rabin, D. and Moore, R. (1984). Heating the Sun's lower transition region with fine-scale electric currents. *Astrophysical Journal*, **285**, 359–367.

Reale, F. and Peres, G. (2000). TRACE-derived temperature and emission measure profiles along long-lived coronal loops: The role of filamentation. *Astrophysical Journal*, **528**, L45–L48.

Roberts, B. and Ulmschneider, P. (1997). Dynamics of flux tubes in the solar atmosphere: Theory. In C.E. Alissandrakis, G. Simnett and L. Vlahos (eds), *Solar and Heliospheric Plasma Physics, Proceedings of the 8th European Meeting on Solar Physics*, Springer, Berlin, pp. 75–101.

Roumeliotis, G. (1991). Joule heating as an explanation for the differential emission measure structure and systematic redshifts in the Sun's lower transition region. *Astrophysical Journal*, **379**, 392–400.

Russell, H.N. (1929). On the composition of the Sun's atmosphere. *Astrophysical Journal*, **70**, 11–82.

Rutten, R.J. and Uitenbroek, H. (1991). Ca II H_{2V} and K_{2V} cell grains. *Solar Physics*, **134**, 15–71.

Schmieder, B. and Mein, N. (1980). Mechanical flux in the solar chromosphere. II. Determination of the mechanical flux. *Astronomy and Astrophysics*, **84**, 99–105.

Schrijver, C.J., Title, A.M., Harvey, K.L., Sheeley, N.R. Jr, Wang, Y.-M., van den Oord, G.H.J., Shine, R.A., Tarbell, T.D. and Hurlburt, N.E. (1998). Large-scale coronal heating by the small-scale magnetic field of the Sun. *Nature*, **394**, 152–154.

Schrijver, C.J. and 16 coauthors (1999). A new view of the solar outer atmosphere by the Transition Region and Coronal Explorer. *Solar Physics*, **187**, 261–302.

Scudder, J.D. (1992). Why all stars should possess circumstellar temperature inversions. *Astrophysical Journal*, **398**, 319–349.

Shoub, E.C. (1983). Invalidity of local thermodynamic equilibrium for electrons in the solar transition region. I. Fokker-Planck results. *Astrophysical Journal*, **266**, 339–369.

Simon, G.W. and Leighton, R.B. (1964). Velocity fields in the solar atmosphere III. Large-scale motions, the chromospheric network, and magnetic fields. *Astrophysical Journal*, **140**, 1120–1147.

Solanki, S.K. (1997). Dynamics of flux tubes in the solar atmosphere: Observations. In G. Simnett, C.E. Alissandrakis and L. Vlahos (eds), *Solar and Heliospheric Plasma Physics*, Lecture Notes in Physics, Springer-Verlag, Heidelberg, pp. 49–73.

Solanki, S.K. (1998). Structure of the solar photosphere. In C. Fröhlich, M.C.E. Huber, S.K. Solanki, R. von Steiger (eds), *Solar Composition and its Evolution – From Core to Corona*, Kluwer, Dordrecht. [*Space Science Reviews*, **85**, 175–186.]

Solanki, S.K., Finsterle, W., Rüedi, I. and Livingston, W. (1999). Expansion of solar magnetic flux tubes large and small. *Astronomy and Astrophysics*, **347**, L27–L30.

Solanki, S.K. and Fligge, M. (2000). Reconstruction of past solar irradiance. In E. Friis-Christensen, C. Fröhlich, J. D. Haigh, M. Schüssler,

R. von Steiger (eds), *Solar Variability and Climate*, Kluwer, Dordrecht. [*Space Science Reviews*, **94**, 139–144.]

Solanki, S.K., Livingston, W. and Ayres, T.R. (1994). New light on the heart of darkness of the solar chromosphere. *Science*, **263**, 64–66.

Solanki, S.K. and Steiner, O. (1990). How magnetic is the solar chromosphere? *Astronomy and Astrophysics*, **234**, 519–529.

St. John, C.E. (1913). Radial motion in sun-spots. I. The distribution of velocities in the solar vortex. *Astrophysical Journal*, **37**, 322–353.

Steffen, M. and Muchmore, D. (1988). Can granular fluctuations in the solar photosphere produce temperature inhomogeneities at the height of the temperature minimum? *Astronomy and Astrophysics*, **193**, 281–290.

Stucki, K., Solanki, S.K., Rüedi, I., Stenflo, J.O., Brković, A., Schühle, U., Wilhelm, K. and Huber, M.C.E. (1999). Coronal holes versus normal quiet sun observed with SUMER. In *VII International Plasma Astrophysics and Space Physics Conference*, Kluwer, Dordrecht. [*Astrophys. Space Sci.*, **264**, 43–52.]

Sturrock, P.A., Dixon, W.W., Klimchuk, J.A. and Antiochos, S.K. (1990). Episodic coronal heating. *Astrophysical Journal*, **356**, L31–L34.

Title, A.M., Tarbell, T.D., Topka, K.P., Fergusen, S.H., Shine, R.A. and the SOUP Team (1989). Statistical properties of solar granulation derived from the SOUP instrument on Spacelab 2. *Astrophysical Journal*, **336**, 475–494.

Tousey, R. (1977). Apollo Telescope Mount of Skylab – An overview. *Applied Optics*, **16**, 825–836.

Ulmschneider, P. (1970). On frequency and strength of shock waves in the solar atmosphere. *Solar Physics*, **12**, 403–415.

Vaiana, G.S. and Rosner, R. (1978). Recent advances in coronal physics. *Annual Reviews of Astronomy and Astrophysics*, **16**, 393–428.

Vernazza, J.E., Avrett, E.H. and Loeser, R. (1973). Structure of the solar chromosphere. Basic computations and summary of the results. *Astrophysical Journal*, **184**, 605–632.

Vernazza, J.E., Avrett, E.H. and Loeser, R. (1981). Structure of the solar chromosphere III. Models of the EUV brightness components of the quiet Sun. *Astrophysical Journal Supplement Series*, **45**, 635–725.

Wang, Y.-M. (1994). Two types of slow solar wind. *Astrophysical Journal*, **437**, L67–L70.

Wildt, R. (1939). Negative ions of hydrogen and the opacity of stellar atmospheres. *Astrophysical Journal*, **90**, 611–620.

Wilhelm, K., Dammasch, I.E., Marsch, E. and Hassler, D.M. (2000). On the source regions of the fast solar wind in polar coronal holes. *Astronomy and Astrophysics*, **353**, 749–756.

Wilhelm, K., Marsch, E., Dwivedi, B.N., Hassler, D.M., Lemaire, P., Gabriel, A.H. and Huber, M.C.E. (1998). The solar corona above coronal holes as seen by SUMER on SOHO. *Astrophysical Journal*, **500**, 1023–1038.

Wikstøl, Ø., Hansteen, V.H., Carlsson, M. and Judge, P.G. (2000). Chromospheric and transition region internetwork oscillations: A signature of upward propagating waves. *Astrophysical Journal*, **351**, 1150–1160.

Wikstøl, Ø., Judge, P.G. and Hansteen, V. (1998). On inferring the properties of dynamic plasmas from their emitted spectra: The case of the solar transition region. *Astrophysical Journal*, **501**, 895–910.

Willson, R.C. and Hudson, H.S. (1991). The Sun's luminosity over a complete solar cycle. *Nature*, **351**, 42–44.

Woods, D.T., Holzer, T.E. and MacGregor, K.B. (1990). Lower solar chromosphere–corona transition region. II – Wave pressure effects for a specific form of the heating function. *Astrophysical Journal Supplement Series*, **73**, 489–512.