Astron Astrophys Rev (1989) 1:3–48

THE
ASTRONOMY
AND
ASTROPHYSICS
REVIEW

Small-scale magnetic structures on the Sun

J.O. Stenflo*

High Altitude Observatory, National Center for Atmospheric Research**, P.O. Box 3000, Boulder, CO 80307, USA

Summary. The Sun provides us with a unique astrophysics laboratory for exploring the fundamental processes of interaction between a turbulent, gravitationally stratified plasma and magnetic fields. Although the magnetic structures and their evolution can be observed in considerable detail through the use of the Zeeman effect in photospheric spectral lines, a major obstacle has been that all magnetic structures on the Sun, excluding sunspots, are smaller than what can be resolved by present-day instruments. This has led to the development of indirect, spectral techniques (combinations of two or more polarized spectral lines), which overcome the resolution obstacle and have revealed unexpected properties of the small-scale magnetic structures. Indirect empirical and theoretical estimates of the sizes of the flux elements indicate that they may be within reach of planned new telescopes, and that we are on the verge of a unified understanding of the diverse phenomena of solar and stellar activity.

In the present review we describe the observational properties of the small-scale field structures (while indicating the diagnostic methods used), and relate these properties to the theoretical concepts of formation, equilibrium structure, and origin of the surface magnetic flux.

Key words: solar magnetic fields – Zeeman effect – plasma physics – stellar activity

1. Introduction

Since Hale's (1908) discovery of magnetic fields in sunspots through the Zeeman splitting of spectral lines, solar magnetic fields have been explored with increasingly sophisticated techniques, revealing the profound and pervasive role of magnetic fields for practically all aspects of solar activity and variability. The sun is unique in being the only star on which the physical processes can be

* On leave from Institute of Astronomy, ETH-Zentrum, CH-8092 Zürich, Switzerland
** The National Center for Atmospheric Research is sponsored by the National Science Foundation

resolved in great detail. It therefore serves as an astronomical reference object, and has sometimes been called the "Rosetta stone" of astrophysics. The sun also provides us with a unique plasma physics laboratory, where a domain of physics can be explored that is not accessible to laboratory experiments, since a scaling of the sun to laboratory dimensions on earth leads to scaled values for the magnetic fields and electric currents that are far beyond what can be reproduced on earth (Alfvén and Fälthammar, 1963).

With improvements in the angular resolution of the observations, but in particular through the use of indirect diagnostic methods that go beyond the resolution obstacle, it has been found that the magnetic field of the sun is extremely structured. Since the size of the basic flux elements is smaller than what can be resolved by existing telescopes, indirect, resolution-independent methods have been developed, showing that more than 90 % of the flux occurs in strong-field form, with the kG flux fragments or "fluxtubes" occupying on the average only a fraction (a few tenths) of one percent of the solar surface. These spatially highly intermittent flux fragments can be found all over the sun, at all heliographic latitudes, and they organize themselves in large-scale polarity patterns, which evolve in a remarkably orderly fashion with the 22 yr magnetic cycle.

The key to a unified understanding of the diversity of phenomena observed on larger scales lies in the understanding of the physical processes at the smallest scales. In the case of the plasma processes in the solar photosphere the smallest relevant length scales are determined by parameters like the pressure scale height (the gravitational stratification of the atmosphere), the photon mean free path (opacity of the plasma), or the length and time scales of ohmic diffusion. Although these scales are smaller than the presently attainable spatial resolution, future telescopes (like the international LEST) may partially resolve them, which is of crucial importance for our understanding of stellar physics.

The sun's small-scale magnetic field is being explored with a variety of methods. Direct mapping of the Zeeman-effect circular polarization recorded in the wing of a Zeeman-sensitive spectral line gives a good qualitative picture of the large-scale distribution, morphology, and evolution of the line-of-sight component of the magnetic flux. Such circular-polarization maps are called magnetograms. However, since the basic fluxtubes are far smaller in size than what has been resolved, the field strengths are not determined by this method, and the obtained values of the magnetic fluxes are subject to large and time-varying errors since the heating processes (affecting the strength of the spectral line used) inside the fluxtubes are not measured. These problems are overcome by using multi-wavelength information (polarized line profiles, combinations of spectral lines), which provide field strengths and temperatures, as well as information on the dynamic structure inside the spatially unresolved fluxtubes, but at the expense of losing information on morphology and evolution. One great advantage of these spectral methods is that they can be directly applied to other stars, since they are resolution-independent. Finally, theoretical MHD models of solar fluxtubes are being constructed at various levels of sophistication, and are being coupled to the

diagnostic problem of inverting polarized spectra. Such computations are very demanding and generally require the use of the fastest supercomputers available.

A large number of reviews dealing with small-scale solar magnetic fields have been published in the past. The empirical aspects have been emphasized in the reviews by Harvey (1977, 1986), Solanki (1987 a, b), Stenflo (1976, 1977, 1978, 1984, 1985, 1986), and Zwaan (1978, 1985, 1987), whereas the theoretical aspects have been emphasized by Parker (1979 a), Nordlund (1983, 1986), Spruit (1983), Spruit and Roberts (1983), and Schüssler (1986, 1987). The interested reader may consult them for further references. The present review will focus on the observational properties of the small-scale magnetic structures (excluding sunspots, which will only be discussed briefly), and relate these properties to the theoretical concepts of formation, equilibrium structure, and origin.

2. Overall organization and evolution of the Sun's magnetic field

2.1. Large-scale polarity patterns

Fig. 1 shows a full-disk magnetogram (circular-polarization map) from NSO/Kitt Peak, recorded on March 7, 1979, during a period of high solar activity. Heliographic north is at top, east to the left. Brighter areas represent line-of-sight magnetic flux directed towards us (positive polarity), darker areas flux directed away from us (negative polarity). Most of the large-scale flux is concentrated in the active latitude belts of the northern and southern hemispheres, with the positive and negative polarity patterns largely oriented in the east-west direction. Young, newly emerged sunspot groups represent the most compact, bipolar magnetic regions. The remnants of old decaying active regions occur where the flux is more widely dispersed and merges into the background field pattern. The correlation time scale for the pattern of a given latitude zone is of the order of a few months, but the flux may actually be replenished over a much shorter time scale (cf. Sect. 5 below).

In the northern hemisphere of the magnetogram in Fig. 1 the compact, bipolar magnetic regions have their polarities oriented negative-positive in the east-west direction, and the polar cap has positive polarity (which is only very weakly discernible in the magnetogram, since the fields close to the limb have their main component perpendicular to the line of sight, and are therefore not well visible in circular polarization). In the southern hemisphere the polarities are reversed. During the previous and following 11 yr activity cycles the polarity situation is the opposite, implying that the full magnetic cycle is 22 yr.

2.2. Field-line topology

A magnetogram like that of Fig. 1 shows the distribution of the *footpoints* of the field lines in the solar photosphere, but it does not tell how the field lines are connected above or below. Since the photospheric spectral lines are most suitable for the Zeeman-effect analysis, in contrast to lines formed higher in the atmosphere, we do not have good magnetic-field observations for the

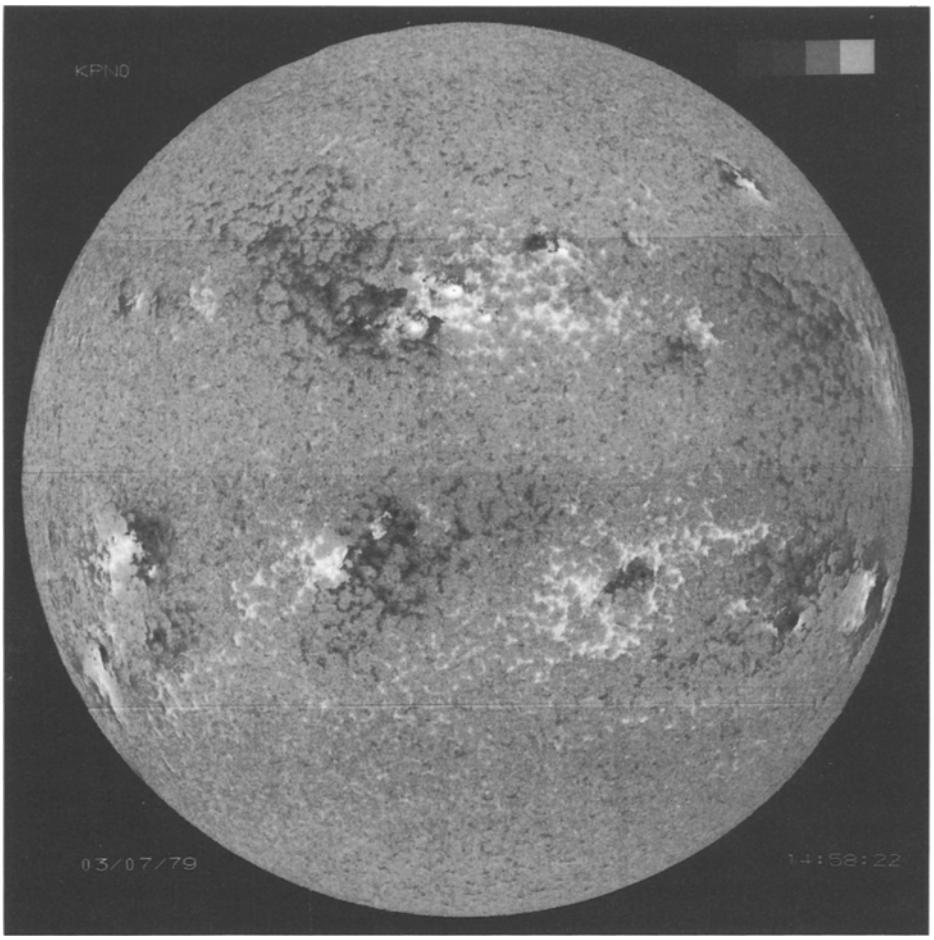

Fig. 1. Full-disk solar magnetogram (map of the line-of-sight component of the magnetic flux) obtained on March 7, 1979, at the National Solar Observatory (Kitt Peak). Brighter areas represent magnetic flux directed towards us (positive polarity), darker areas flux directed away from us (negative polarity). (Courtesy of J.W. Harvey, National Solar Observatory.)

chromosphere or corona. Soft X-ray images of the sun however show loop-like emission structures in projection on the solar disk, outlining regions of enhanced density and temperature bridging magnetic regions of opposite polarity. Using the coronal X-ray emission as a "proxy" for the magnetic field lines, the topological field connections can be seen. The strongest connections occur between the neighbouring polarities in active regions, but surprising connections across the equator between active regions in the north and south hemisphere are also found, showing that the fluxes of opposite polarities do not fully balance each other within a given active region.

From the point of view of their coronal effects, the magnetic fields can be divided into two main topological classes: *open* and *closed* fields. In our

terminology a field line is "open" if it is pulled out by the expanding corona (solar wind) into interplanetary space. Field lines that loop back to the solar surface before reaching heights (at most a couple of solar radii above the solar surface) at which the solar wind can pull them out in the radial direction are "closed".

This distinction between open and closed is of profound importance for the temperature-density structure of the solar corona. In open-field regions the corona can expand unimpeded into interplanetary space. In such regions the coronal density is low, leading to much reduced X-ray emission. From their appearance on X-ray photographs, such regions have been called "coronal holes". In closed, loop-like regions, on the other hand, the plasma and the heat input are "bottled up", without a "safety valve". The coronal plasma in the closed magnetic regions has much higher densities and temperatures as compared with their surroundings, making these regions high-contrast emission features in X-ray images.

We can further distinguish between two different types of closed magnetic regions, at least with respect to their coronal effects: (1) Loops in ordinary active regions. (2) X-ray bright points (XBP), which coincide with small, bipolar magnetic regions, and are ubiquitous all over the sun. The XBPs have short lifetimes, typically 8–12 hr.

As it has not been possible to directly measure the magnetic field above the photosphere (apart from very indirect or fragmentary information from proxies, coronal emission-line polarization, or radio data, cf. Stenflo, 1978), methods to compute the coronal field by solving Maxwell's equations for a vacuum field, using the observed line-of-sight component of the photospheric magnetic flux as an inner boundary condition, have been employed. To mathematically mimic the effect of the solar wind, an outer spherical boundary or "source surface" has been introduced, at which the field is postulated to be oriented in the radial direction (e.g. Altschuler and Newkirk, 1969).

Although qualitative agreement between the shape of large-scale coronal emission structures and the computed coronal field has been found, there are a number of reasons why such extrapolations of the measured photospheric fields should be used with caution: (1) The measured photospheric magnetic fluxes may be in error (the typical order of magnitude of the error being a factor of two) due to fluxtube heating processes or other spatially unresolved effects. (2) The solar photosphere cannot be observed from all sides at once, so the inner boundary values have to be collected over the course of one solar rotation (about 27 days), during which time the field may undergo large evolutionary changes. (3) The distance from the sun at which the field becomes radial depends on the field strength and topology as well as on the coronal temperature and density. The true "source surface" will therefore be far from spherical. (4) The full vector magnetic field at the photosphere is not observed, only the line-of-sight component. The usual assumption that the field direction is determined by the solution for the outside vacuum field, using the observed line-of-sight component as a constraint, is certainly not correct, since the field is far from force-free (and thus also far from vacuum) where it is measured (in the photosphere). The orientation of the discrete, kG fluxtubes is instead dominated by local forces,

in particular by the buoyancy force induced by the steep density stratification, without the fluxtubes "feeling" much of the magnetic tension related to the global, topological connections in the corona. Due to the importance of buoyancy, a better assumption for this boundary condition is that the average field direction is radial (Stenflo and Vogel, 1986), although large local deviations from the radial direction may occur, in particular in young bipolar magnetic regions. (5) The vacuum-field assumption ignores pressure forces or the existence of coronal electric currents, which are generally present, in particular above active regions.

2.3. Size distribution

It may be noted in Fig. 1 that the background magnetic fluxes are not smoothly distributed, but occur as flux fragments forming a "network" pattern with a characteristic cell diameter of about 30,000 km. The magnetic network coincides with the emission network seen in different spectral lines, with highest contrast for lines formed in the chromosphere-corona transition region, where the heating of the corona takes place. This indicates that the magnetic fields play an important role for the heating and energy balance of the outer solar atmosphere. The network also coincides with the boundaries of the photospheric velocity cells that constitute the so-called supergranulation. The largely horizontal motions of the plasma inside the supergranular cells sweep the "frozen-in" field lines towards the cell boundaries.

The magnetic flux is further broken up in smaller fragments by turbulent eddies of smaller sizes. Below the supergranulation scale comes the mesogranulation, but more conspicuously the still smaller granulation, with typical cell diameters of 1000 km and autocorrelation lifetimes of 7 min (in contrast to 24 hr for the supergranulation). During exceptional seeing conditions it has been possible to see how the tiny flux fragments are preferentially located at the granular cell boundaries (Title et al., 1987), the intergranular lanes, where the velocity field becomes vertical (downdrafts). This preference for the intergranular lanes can more easily be seen when observing the morphology of the emission structures or small bright points as proxies for the magnetic-field concentrations (e.g. Mehltretter, 1974). Spruit and Zwaan (1981) measured the size distribution of continuum bright points in a young active region and used it as being representative of the fluxtube size distribution. The observed sizes of the emission features may however not represent the fluxtube sizes, as explained in Sect. 3.6.

The flux elements are neutrally unstable with respect to shape deformations, and are subject to fragmentation by the interchange or flute instability. Extensive numerical model calculations by Nordlund (1983, 1986) of the 3-dimensional structure and evolution of solar granulation with magnetic fields, solving the time-dependent MHD equations with proper density stratification and radiative energy transport, show how the flux elements are deformed, fragment, and coalesce with each other, with the evolution of the granulation pattern. Although the flux is not destroyed, the flux elements have a limited lifetime (comparable to that of the turbulent eddies) as individual entities due to the morphological changes associated with reshuffling of the field lines (which will lead to many

field line crossings or topological "catastrophes" in the field geometry, at which rapid reconnection can take place).

The numerical resolution of these 3-D simulations is however not adequate for resolving the internal structure of the flux elements, simulating the process of "convective collapse" (cf. Sect. 5.1), or treating the effect of turbulent eddies substantially smaller than the solar granulation (these small-scale eddies are inferred from the observed "microturbulent" Doppler broadening of spectral lines). These simulations therefore do not tell us what happens at the small-scale end of the spectrum of sizes. Observations, however, provide us with some powerful constraints, as explained next.

In general magnetic flux appears to emerge from the solar interior in the form of larger, coherent fluxtubes (the largest ones being the sunspots), which decay by fragmentation at a rate of about 10^{15} Mx s^{-1}, thus transferring the flux from smaller to larger spatial wavenumbers in the Fourier domain (Gokhale and Zwaan, 1972; Harvey and Harvey, 1973; Stenflo, 1976). The size distribution of sunspot umbrae is found to be lognormal, consistent with a fragmentation origin (Bogdan et al., 1988). It is however quite remarkable that the field does not diffuse very much in the process of fragmentation, but that the strong, kG fields are retained, although the flux fragments become more and more separated by almost field-free areas. This property can be observed directly for fragments down to the spatial resolution limit (about one arcsec, which equals 725 km), but it can also be determined beyond the resolution limit. As will be explained in Sect. 3, it can be shown that more than 90 % of the total magnetic flux penetrating the photosphere outside sunspots occurs in kG form (field strengths of 1–2 kG in the photospheric layers, as compared with typically 2–4 kG in larger sunspots). It can further be shown that all these non-sunspot kG flux elements have sizes smaller than one arcsec, generally much smaller (a typical estimate being 100 km), thus in most cases being well beyond the resolution limit of exisiting instruments.

If the flux fragmentation were allowed to proceed indefinitely to smaller sizes, a scale (in the km range) would eventually be reached at which it would no longer be possible to contain the kG fields as embedded in an almost field-free environment, due to the efficiency of ohmic diffusion at these scales. It would thus seem that fragmentation, followed by ohmic diffusion, would make the flux end up in a diffused, weak-field state. The fact that more than 90 % of the flux is in kG form shows however that most of the flux never ends up in a diffused state. This suggests that at some scale well above the ohmic diffusion scale (but much smaller than the granulation scale) the rate of flux fragmentation is counteracted and balanced by the rate of flux concentration and coalescence. This scale then constitutes the effective end of the size spectrum of magnetic elements, at which most of the flux will pile up. In Sect. 5.1 we will describe a mechanism by which diffuse flux spontaneously collapses into a strong-field (kG) state. This convective instability loses its efficiency for fluxtube sizes smaller than about 100 km, due to the efficient thermal exchange between fluxtube interior and exterior for optically thin fluxtubes. It is thus natural to expect that the fluxtube size spectrum should end at scales of the order of 50–100 km, at least not very far below that. In Sect. 3.6 we will present some observational evidence, due to Zayer et al. (1989), that

the magnetic flux elements outside sunspots really have sizes of this order, even when the magnetic filling factor is relatively large (in a strong plage).

One issue that we have not discussed so far is the state of the weaker fields between the isolated kG flux fragments. This difficult question, including the possibility that the magnetic polarities are mixed on scales smaller than the resolution limit ("turbulent" magnetic field), will be addressed in Sect. 4.

2.4. Small-scale evolution

Many different physical processes affect the evolution of the magnetic fluxes that we observe in the solar photosphere. The magnetic field is believed to be generated in the solar interior by dynamo processes, and brought to the surface by buoyancy forces in the convection zone. Convective eddies carry the "frozen-in" flux elements around in a random-walk process (cf. Simon et al., 1988), leading to what has been called "turbulent diffusion" of the flux, both within the convection zone and on the solar surface. The constant supply of new flux from the sun's interior during the course of the activity cycle has to be balanced by the rate of flux removal, otherwise there would be an excessive accumulation of flux at the solar surface. The main candidate mechanisms for flux removal are (1) submergence, whereby flux loops are being pulled down below the surface, and (2) annihilation of encountering opposite-polarity fluxes by ohmic diffusion, leading to field-line reconnection.

The processes of emergence, submergence, and annihilation are poorly understood, since most of them occur rapidly on a small scale beyond what can be resolved. As the magnetic fluxes outside active regions are small, time integrations of minutes have to be used to obtain sufficient polarimetric accuracy due to the limited photon-collecting areas of current instruments, while the atmospheric seeing (due to turbulent fluctuations in the earth's atmosphere) varies mainly on time scales shorter than 20 ms. Therefore few good magnetogram sequences showing the evolution of small-scale fields are available.

Fig. 2 illustrates one sequence recorded on June 22, 1985, in a quiet region near the center of the solar disk with the video magnetograph of the Big Bear Solar Observatory (Wang, 1988). The sequence spans 10.5 hr (the universal time is indicated at the lower left hand corner of each of the 8 frames). As usual, the opposite polarities are represented by brighter and darker areas, and as is typical for quiet regions, the magnetograms have a "salt and pepper" appearance due to the mixing of opposite-polarity flux fragments over the field of view. The flux elements in the frames appear as discrete clumps separated by apparently field-free areas, but the elements are far from spatially resolved. The field of view is about $300 \times 200\,\mathrm{arcsec}^2$, and the length of 10,000 km is indicated at the lower right corner of the figure. The peak flux density directly observed in the stronger elements of the sequence was 100–200 G. As it is known via indirect diagnostic methods (cf. Sect. 3) that the actual field strengths are 1–2 kG, and that the characteristic fluxtube sizes are rather in the range around 100 km, the flux "clumps" that we see in Fig. 2 must consist of collections or clusters of smaller, unresolved flux elements, with magnetic filling factors (fraction of the

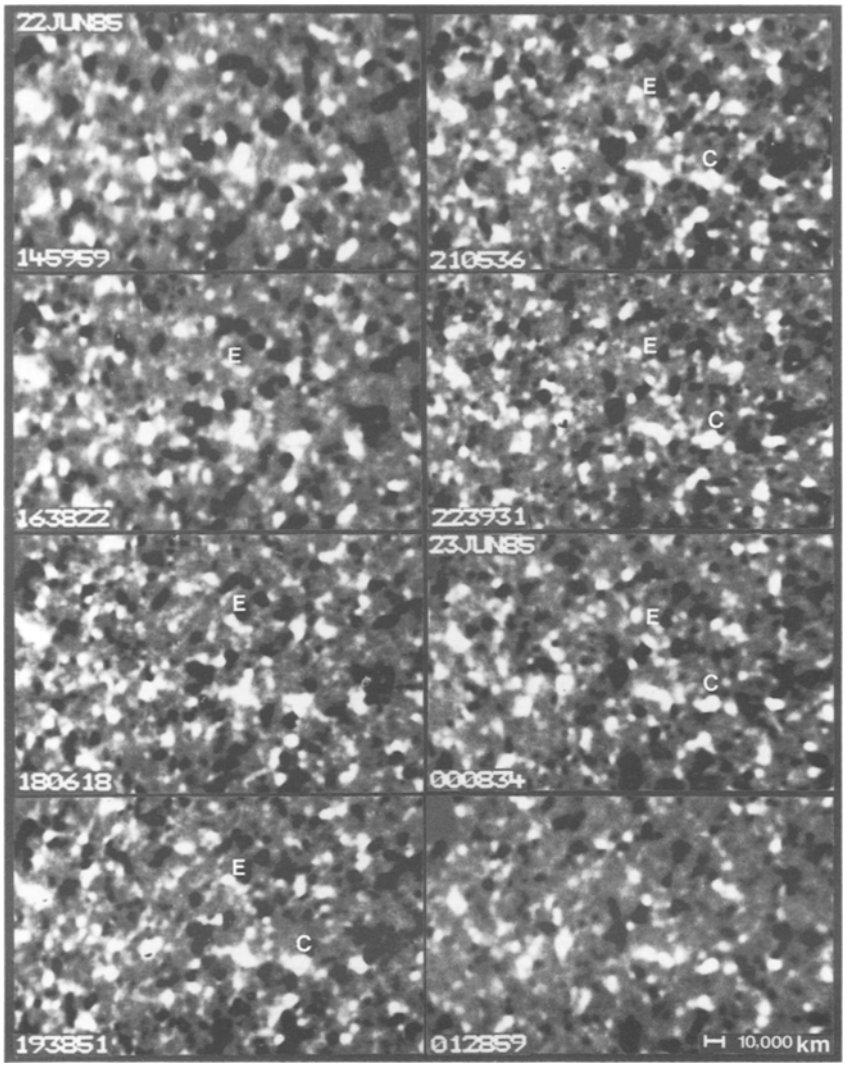

Fig. 2. Sequence of 8 magnetograms, showing the evolution of the discrete flux elements in a quiet region of mixed magnetic polarities at disk center, recorded on June 22-23, 1985, at the Big Bear Solar Observatory (Wang, 1988). The universal time is indicated at lower left in each frame. The field of view is about 300×200 arcsec^2 (the scale is indicated at the bottom right of the figure). The letters "E" and "C" indicate an emergence and a cancellation feature, respectively. (Courtesy of H. Wang, BBSO.)

surface area occupied by the fluxtubes inside the spatial resolution element) of about 10 % or less within a flux "clump".

The evolution of such a complicated pattern can best be followed when the frames are displayed as a movie, which shows the flow patterns through the proper motions of the observed structures. Two features have been specially marked in Fig. 2: "E" denotes a small emerging bipolar region, while "C" denotes a

"cancellation feature", a small bipolar region for which the opposite polarities move together and cancel each other. Many such features can be observed in similar sequences (Livi et al., 1985; Martin et al., 1985 a, b). It has however not been possible to conclude from the observations to what extent the observed instances of flux cancellation are due to flux submergence or reconnection (flux annihilation). Observations with much higher spatial and time resolution are needed for this purpose.

The preceding discussion has dealt with the qualitative evolutionary changes. When attempting a *quantitative* interpretation of the magnetograms, however, great care has to be exercised. Observations by Wallenhorst and Howard (1982) and by Wallenhorst and Topka (1982) that the apparent line-of-sight magnetic flux in active regions changed by typically 30–40 % when the sunspots disappeared, without any observed spreading of the flux or cancellation with opposite-polarity flux, led to claims that some mysterious flux removal mechanism was operating. It is however all too often forgotten that "magnetograms" are simply maps of the circular polarization, not of the magnetic flux, although the close connection between circular polarization and line-of-sight magnetic flux has led to the term "magnetogram". The relation between polarization and magnetic flux changes with the internal temperature and opacity structure of the fluxtubes. Highly fragmented flux is hotter (cf. Sect. 3.4) than less fragmented flux. It therefore generally produces less polarization, due to the more pronounced weakening of the spectral lines. As the flux evolves by fragmentation, the integrated amount of circular polarization will thus decrease, although the integrated flux is conserved. This provides a natural explanation of the above-mentioned observations (Stenflo, 1984; Grossmann-Doerth et al., 1987).

2.5. Active regions and sunspots

Active regions represent the largest concentrations of magnetic flux on the sun. They emerge from the solar interior as compact, bipolar magnetic regions generally containing sunspots, and decay over a period of months by fragmentation and spreading of the flux fragments over a wider area by turbulent diffusion.

Fig. 3 illustrates the magnetic fields in a young active region recorded on August 6, 1987, at NSO/Sacramento Peak (Tarbell et al., 1988; Topka et al., 1988). To the left is a magnetogram, to the right the chromospheric emission in the red wing of the Hα line. The field of view is $80 \times 90\,\mathrm{arcsec}^2$, and the tick mark spacing is approximately 2.6 arcsec. Images of the Doppler velocity and of the continuum and line brightness were also recorded.

Although the magnetic polarities are separated on a large scale, they are found to mix and stream past each other on a small scale. The dark Hα filaments indicate largely horizontal fields in the chromosphere, connecting footpoints of opposite magnetic polarities. At the footpoints, or end points of the dark filaments, one often sees bright points. Emerging fluxtubes are first noticed by causing dark striations in the continuum and line brightness, accompanied by upwards motion. This is followed by the appearance of the bright end points, which are accompanied by concentrated downflows.

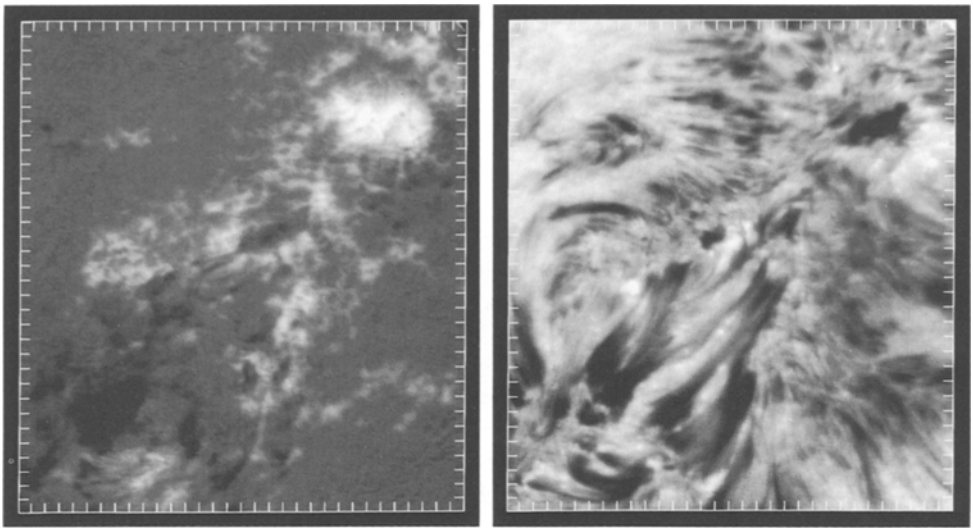

Fig. 3. Magnetogram (left) and Hα red wing filtergram (right) of an active region, recorded by the Lockheed SOUP instrument at NSO/Sacramento Peak on August 6, 1987 (Tarbell et al., 1988; Topka et al., 1988). The field of view is about 80×90 arcsec^2, the tick-mark spacing approximately 2.6 arcsec. (Courtesy of Z.A. Frank, T. Tarbell, and A. Title, Lockheed Palo Alto Research Laboratories.)

Sunspots, which are part of active regions, represent the largest "fluxtubes" with the strongest fields. In white-light photographs they are characterized by the dark umbra (with a continuum brightness typically about 5 % that of the undisturbed photosphere), which often is filled with bright "umbral dots", and the filamented penumbra. Fig. 4 shows two recent recordings (Lites and Scharmer, 1988) from the Swedish La Palma observatory. To the left are white-light slit-jaw images, showing the location of the spectrograph slit. To the right the Zeeman splitting, Doppler shifts, and brightness variations of the Fe I 6302 Å line are displayed. The two surrounding spectral lines are telluric lines, and are therefore narrow and straight. The upper recording in Fig. 4 shows the Zeeman splitting and continuum fluctuations in the penumbra, while the lower recording also includes spectra of the umbra and of umbral bright dots. Note in particular how in the upper part of the spectrum at the bottom of Fig. 4 the Zeeman splitting suddenly vanishes within an extremely narrow (subarcsec) region, indicating large magnetic-field gradients (which suggests strong electric currents). The quantitative analysis of these recordings has however not yet been completed.

One fundamental question concerning the sunspots is the inclination of the magnetic field in the penumbra. Polarimetric observations suggest that the inclination with respect to the vertical direction varies smoothly from zero at the center of the umbra to $70 - 90°$ at the outer edge of the penumbra, with a substantial vertical field component over most of the penumbra (Deubner and Göhring, 1970; Wittmann, 1974; Lites and Skumanich, 1989), while other data (Giovanelli, 1982) and theoretical penumbra models tend to favor an almost

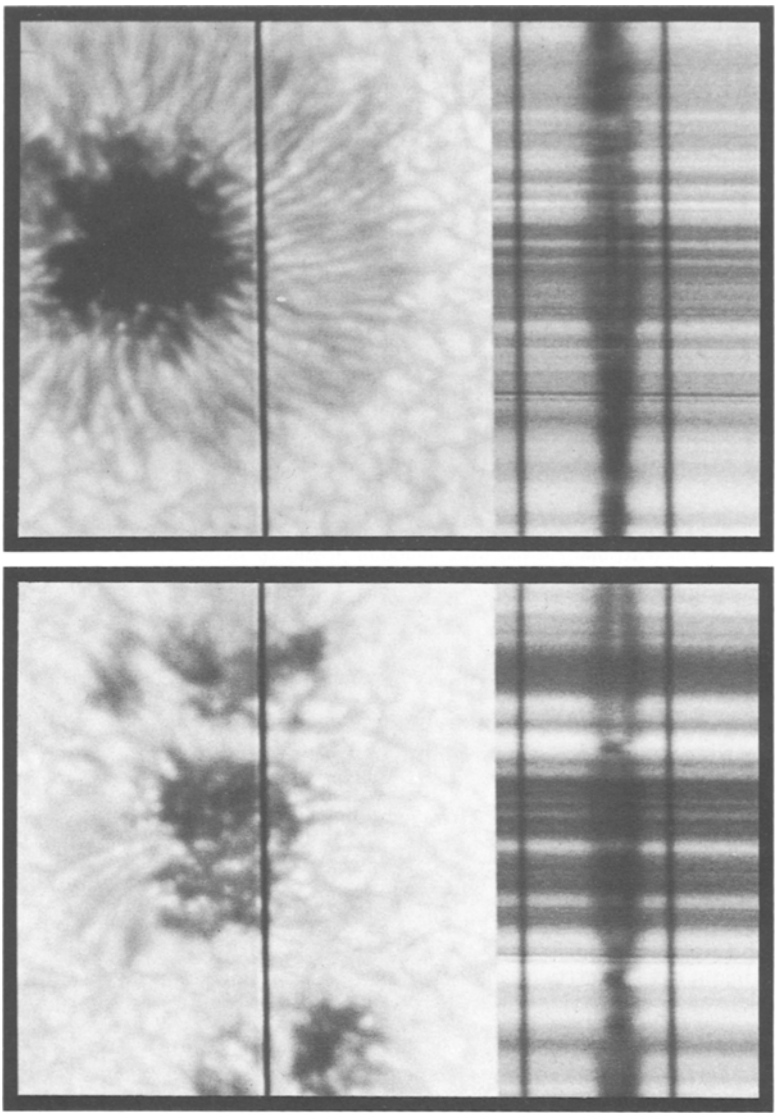

Fig. 4. Slit-jaw white-light image (left) and spectrogram (right) of two sunspots, one with a prominent penumbra (top), the other with numerous bright dots in the umbra (bottom), obtained by Lites and Scharmer (1988) on October 29, 1987, at the Swedish La Palma observatory. The spectra show the Zeeman splitting in the Fe I 6302 Å line, which is surrounded by two narrow telluric lines. The field of view along the slit is 26 arcsec. (Courtesy of B.W. Lites, High Altitude Observatory.)

purely horizontal field throughout the penumbra. Due to the large influence of straylight in the penumbra the observational picture is not yet well established, so this remains an open issue.

Another basic question concerns the stability of sunspots and the nature of the umbral bright dots. Parker (1979 b) and Spruit (1981) have argued that the sunspot will split up in smaller flux ropes immediately below the solar surface

due to the interchange instability, which is suppressed in the layers above due to the stabilizing effect of the buoyancy force. Field-free gas between the separated flux ropes may penetrate into the visible layers, appearing as bright dots in the umbra. This picture could be tested by determining whether the magnetic field in the umbral dots is much weaker than in the dark umbra. Such a determination is however also impeded by stray light problems, so no conclusive results have yet been obtained. One way to get around this is to record the Zeeman splitting or polarization in the spatially unresolved umbra simultaneously in two spectral lines of greatly different temperature sensitivities, and apply a two-component model for the interpretation. An approach of this type (but without polarization optics) has been employed by Kneer (1973), using the lines Fe II 6149.2 Å (for the umbral dots) and Ca I 6156.0 Å (for the dark umbra). He found the field strength to be a factor of two smaller in the umbral dots as compared with the surroundings, in contrast to later work by Zwaan et al. (1985) and Lites and Scharmer (1988), who found no great difference in field strength. This issue however remains unsettled.

A new diagnostic method of great promise that has recently been applied to sunspots (Deming et al., 1988) is the use of the infrared Mg I 12.32 μm line (cf. also Brault and Noyes, 1983; Zirin and Popp, 1989). Due to its large wavelength it has a Zeeman splitting much larger than other lines that have been used, and it appears to be formed in the low chromosphere, where few other suitable lines for sunspot diagnostics are available.

In a unified view, sunspots represent the large-scale portion of the size spectrum of fluxtubes. Empirically we find that fluxtubes with a diameter larger than about one arcsec (725 km) appear as darker photospheric structures, whereas smaller fluxtubes appear as brighter features. This can be given a theoretical explanation, as we will see in Sect. 5.

3. Observational properties of magnetic fluxtubes in the solar photosphere

3.1. Diagnostic methods

Magnetograms represent maps of the amount of circular polarization in the wing of a selected spectral line. In a first approximation this measured quantity is proportional to the amount of magnetic flux along the line-of-sight direction through the spatial resolution element (smeared by atmospheric seeing) of the instrument. As however the intrinsic field strength in the flux elements is in the kG range and does not vary much over the sun (as will be explained in Sect. 3.2), the quantity displayed in magnetograms can also be looked upon as the magnetic filling factor, i.e., the fractional surface area covered by the dominating sources of magnetic flux (the kG "fluxtubes").

The measured flux values or filling factors may however be greatly in error, because the relation between the amount of circular polarization and line-of-sight magnetic flux is generally calibrated using the *average* line profile or solar model atmosphere, which differ greatly from the corresponding profile or atmosphere inside the magnetic elements, due to the very different physical conditions inside

the magnetic fluxtubes. The problem of magnetograph calibration is that the conditions inside the fluxtubes are not directly accessible, since the fluxtube sizes are much smaller than the resolution limit of the telescopes. Indirect diagnostic methods have to be developed to bypass this problem.

It is thus clear that the problem of an accurate determination of magnetic fluxes is intimately linked to a variety of other fluxtube parameters, like the internal thermodynamic structure, and the combined effects of these factors have to be unraveled from each other. This can be done by using simultaneous polarimetric information at different wavelengths, which actually provides us with a tremendous diagnostic opportunity. As we will see below, the objective of the various diagnostic methods is not at all to merely establish a correct "calibration" of solar magnetographs (this is just a byproduct). Instead the aim is to derive complete empirical fluxtube models, including the height variations of the magnetic field strength, temperature, and mass motions, and even to determine the sizes (in contrast to the filling factors, which only provide upper limits to the sizes) of the spatially unresolved magnetic elements.

The situation is reminiscent of that in stellar spectroscopy. In spite of the circumstance that the stars are spatially unresolved point objects, it has been possible to extract a great deal of information about the stratification of temperatures, densities, velocity fields, or element abundances, using information not at one wavelength only, but through disentangling the combined information stored in many spectral lines. Similarly, in the case of the spatially unresolved solar magnetic fluxtubes, the influence of the various physical parameters and their height variations can be disentangled by combining the polarization information recorded simultaneously in different carefully selected spectral lines. As we deal with spatially unresolved structures, these methods developed for use on the sun will work on other stars as well.

The spatial-resolution limit is thus overcome by using additional information in the spectral domain. The reason why this has worked so exceptionally well for the fluxtube physics is due to a fortunate circumstance: A two-component model, with one magnetic and one non-magnetic component, has turned out to be a good approximation for the sun, and is likely to remain so for solar-type stars as well. The empirical evidence for the applicability of a two-component approach will be discussed further in Sect. 3.3 (cf. also Stenflo, 1976). With more sophisticated methods we can now also study more general cases with deviations from the two-component approach, including continuous spatial variations of the parameters.

In terms of such a two-component model, the amount of circular polarization, or Stokes V (with the full state of polarization being described in terms of the four Stokes parameters I, Q, U, and V, cf. Chandrasekhar, 1950), originates only in the magnetic component, regardless of how small the magnetic filling factor is, while Stokes I (the ordinary intensity) has contributions from both components. The Stokes V spectrum thus allows us to explore the physical conditions inside the magnetic fluxtubes, independent of how small they are, uncontaminated by any "stray light" from the surrounding, non-magnetic atmosphere.

Some of the various spectral diagnostic methods that have been applied are the following: (1) The line-ratio technique, which uses the ratio between the Stokes V signals recorded simultaneously in two or more carefully selected spectral lines. Depending on what parameter is isolated the most by the line selection we speak of for instance a "magnetic" line ratio or a "thermal" line ratio. (2) The shape of the Stokes V profiles or of the wavelength-integrated V profiles (called I_V, cf. Sect. 3.4) is particularly useful for studying the internal dynamics of the fluxtubes. In the infrared the Zeeman splitting can be measured directly, and the Zeeman broadening due to spatial magnetic-field variations in the fluxtube can be separated from the Doppler broadening. (3) The statistical properties of the Stokes V profile parameters for a large number of spectral lines of different line strengths provide powerful information for determinations of the height variations, in particular of the internal fluxtube temperature. (4) The center-to-limb variations of the above parameters provide additional constraints on the height variations as well as on the fluxtube sizes. (5) The above methods applied to the Stokes Q and U (not yet tried for U) parameters provide information on fluxtube tilts as well as additional constraints on the possible empirical models. Since the fluxtube properties vary slightly with the magnetic filling factor, all these methods need to be applied to regions of different filling factors.

In the following subsections we will summarize some of the results achieved with these methods, and at the same time try to indicate how the information can be extracted. For more details we refer to the various reviews on the empirical fluxtube properties mentioned at the end of Sect. 1.

3.2. Field strengths

For weak magnetic fields the circular polarization (Stokes V) is proportional to the field strength, or, if the fields are not spatially resolved, to the amount of flux (along the line of sight) through the spatial resolution element. If, for pedagogical reasons, we consider a two-component model with filling factor α, the average field strength, proportional to the magnetic flux, is αB. We can then write

$$V = k \alpha B , \tag{1}$$

where k in the weak-field limit is a constant that is determined by magnetograph calibration (which however is subject to serious errors due to temperature and field-strength effects, see below). Knowledge of this constant makes it possible to translate polarization to magnetic flux. For simplicity, consider the case of a normal Zeeman triplet and a purely longitudinal magnetic field. The spectral line is then split in two σ components of opposite circular polarization, shifted to higher and lower wavelengths by

$$\Delta\lambda_{\mathrm{H}} = 4.67 \times 10^{-13} \, g \, \lambda^2 B , \tag{2}$$

where g is the Landé factor, the field strength B is given in G, and the wavelengths are in Å. Stokes V can then be written as

$$V = \frac{1}{2} \alpha \left[I_\alpha(\lambda + \Delta\lambda_H) - I_\alpha(\lambda - \Delta\lambda_H) \right] . \tag{3}$$

Here we have written I_α instead of simply I, to emphasize that it is the intensity profile I_α inside the magnetic component that has to be used, not the average intensity profile I, which is the directly observed quantity. For weak fields, we may use the Taylor series expansion

$$V = \alpha \, \Delta\lambda_H \, \frac{\partial I_\alpha}{\partial \lambda} + \ldots , \tag{4}$$

and retain only the first term. Since $I_\alpha \neq I$, mainly due to temperature line weakening inside the flux regions, the factor k in Eq. (1) cannot be determined from Eqs. (2) – (4), unless additional information on the spatially unresolved line weakening is available. Normally, however, the calibration constant used in magnetograph work is based on the (incorrect) assumption that $I_\alpha = I$.

The weak-field assumption means that the Zeeman splitting is assumed to be small in comparison with the line width. If this condition is not satisfied, the higher-order terms in the Taylor series expansion (4) become important, i.e., the relation between V and B becomes non-linear. This non-linearity varies from line to line depending on Landé factor and line width, and it contains information on the field strength B (which determines the Zeeman splitting $\Delta\lambda_H$), in contrast to the flux αB.

The polarization V recorded in any given spectral line is subject to the combined influence of the magnetic flux (filling factor), field strength, and line weakening. To unravel these competing effects the *line-ratio* method was developed (Stenflo, 1968, 1973). By selecting two lines of the same multiplet, having the same line strength and excitation potential, but differing only in their Landé factors, both the filling factor and the line weakening effect will cancel out when forming the ratio between the V signals recorded simultaneously in the two lines. If the magnetic field in the fluxtubes were intrinsically weak ($\lesssim 500$ G), the apparent magnetic fluxes (derived using the weak-field approximation of Eq. (4)) would be identical in the two lines. Any deviation between the apparent fluxes is due to the differential non-linearities (Zeeman saturation) in the two lines, and the value of this deviation allows the intrinsic field strength to be derived.

The two main lines used for this technique have been Fe I $\lambda\lambda$5247.06 and 5250.22 Å, which both belong to multiplet no. 1 of iron, and have effective Landé factors of 2.0 and 3.0, respectively. The method was first used on a quiet region at disk center, and although the filling factors in most of the sampled disk positions were well below one percent (the apparent flux densities were only of the order of a few G), the intrinsic field strengths derived from the line ratio were 1–2 kG (Stenflo, 1973).

An illustration of how the line-ratio technique was applied in the 1970s, using ordinary (Babcock-type) magnetographs recording the circular polarization simultaneously in two spectral lines, is shown in Fig. 5, based on the data used by Frazier and Stenflo (1978). The apparent field strengths in the two lines, B_{5247} and B_{5250}, derived from the magnetograph signals using Eq. (4) with I_α replaced by I, are plotted against each other for each point in the raster scan across the disk (the scan covered a plage area without sunspots, near disk center). If the fields were intrinsically weak, all the points would fall along the 45° line. Instead

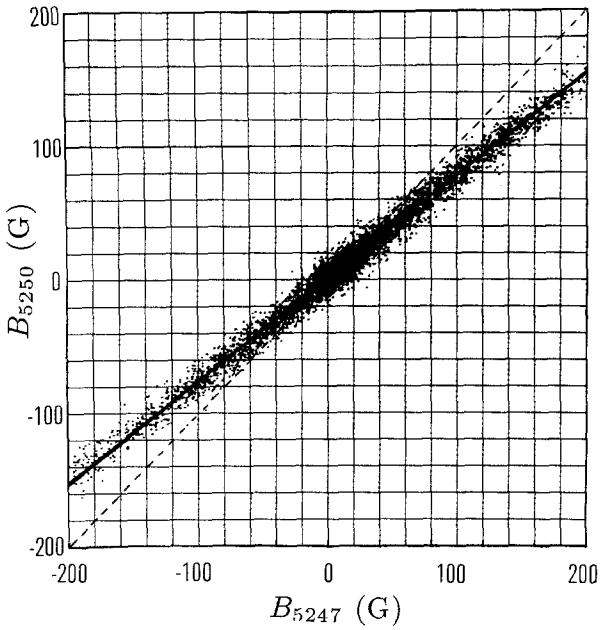

Fig. 5. Scatter plot of the apparent flux densities observed simultaneously in the Fe I 5247.06 (horizontal axis) and 5250.22 Å (vertical axis) lines, based on the data used in Frazier and Stenflo (1978). If the fields were intrinsically weak ($\lesssim$ 500 G), the points would fall around the 45° line (dashed). From the *slope* of the regression line, the intrinsic field strength can be determined

they fall along a well-defined line with a smaller slope. It is the value of the *slope* that gives us the intrinsic field strength in the unresolved flux elements, independent of the apparent B values (which are proportional to the magnetic filling factor).

Notice in Fig. 5 that the slope of the regression relation between B_{5250} and B_{5247} appears to be constant as we move out from the origin, independent of apparent field strength or filling factor. This indicates that the variation of the intrinsic field strength with filling factor is small. A later investigation revealed a filling-factor dependence (Stenflo and Harvey, 1985), but the intrinsic field strength increased only by about 20 % when the filling factor increased by a factor of 6 (when going from the quiet network to a plage). This suggests that the fluxtubes in the quiet network are very similar to those that make up plages, the number density and collective filling factor are simply much larger in plages.

A new breakthrough in the diagnostics of spatially unresolved solar magnetic fluxtubes came with the conversion of the Fourier transform spectrometer (FTS) at the Kitt Peak McMath telescope into a polarimeter. The polarized spectrum can be recorded fully spectrally resolved, with a simultaneous wavelength coverage determined only by the band width of the prefilter used, typically 1000 Å. A 6 Å portion of a Stokes I and V FTS recording in a plage near disk center is shown in Fig. 6. It can be used to illustrate several of the points made above, and

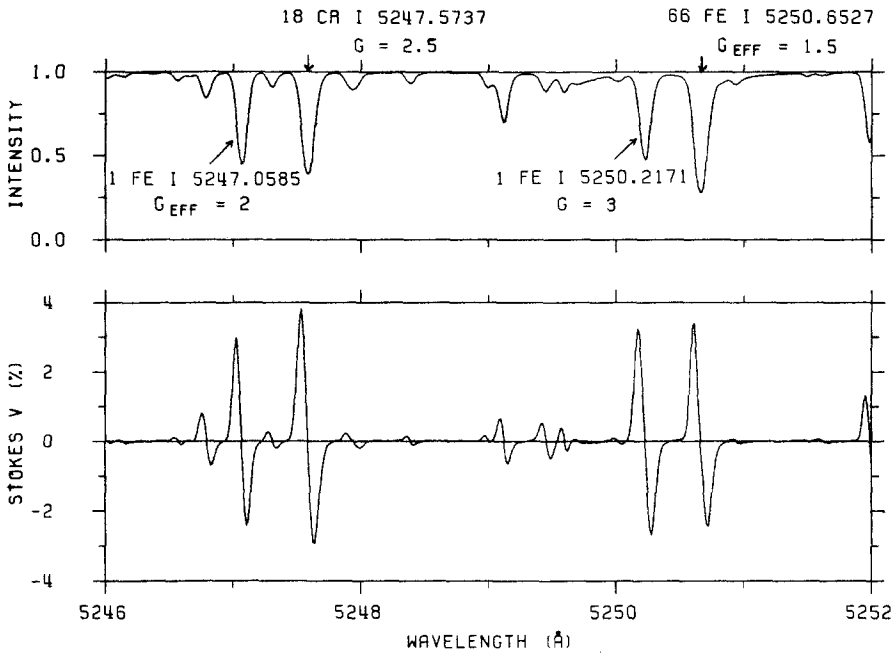

Fig. 6. Intensity and circular polarization (Stokes V) as a function of wavelength, recorded in a plage near disk center with the FTS polarimeter of the NSO McMath telescope (Stenflo et al., 1984)

will also serve to explain the diagnostic methods used for internal temperature and velocity determinations, to be discussed in Sect. 3.4 and 3.5.

In a first approximation the Stokes V spectrum in Fig. 6 reproduces the shape of $\partial I/\partial\lambda$, as expected from Eq. (4). A closer look however reveals a number of important differences, some of which will be mentioned in later sections. It is in these differences that the important diagnostic information is contained. Focusing our attention for the moment on the two Fe I lines at 5247.06 and 5250.22 Å, we would expect their Stokes V amplitudes to be in the proportion 2:3 (of their Landé factors) if the fields were intrinsically weak, but the actual ratio is closer to 1:1. In addition we may discern that the Stokes V profile is broader for the 5250.22 Å line as compared with the 5247.06 Å line. These two effects — Zeeman saturation of the polarization amplitude, and Zeeman broadening of the profile — both result from the non-linearities, or higher-order terms in Eq. (4). These higher order terms are much more important for the 5250.22 Å line, which has the larger Landé factor. With the FTS observations, we can follow such effects throughout the line profiles of many hundreds of simultaneously recorded spectral lines, in contrast to magnetograph recordings represented by Fig. 5, which only sample one wavelength window in each line. The magnetograph recordings on the other hand allow us to sample a large number (thousands) of spatial points, in contrast to the FTS, which typically requires 1/2 h for each spatial point. There is thus an extreme tradeoff between spatial and spectral information in these two methods, and they complement each other.

3.3. Field-strength distributions, and amount of flux in kG form

Although the line-ratio method reveals the presence of kG fields in an unambiguous way and leads to a good estimate of the field strength, a detailed quantitative interpretation, e.g. of the slope in the scatter plot of Fig. 5, is model dependent for the following reasons: (1) The unresolved magnetic elements within the spatial resolution element may not be characterized by one field strength only, but there may be a distribution of field strengths, both horizontally and vertically. (2) The values of the derived field strengths are sensitive to the amount of Doppler broadening assumed (which is generally larger than in the external atmosphere), since the non-linearities scale with the ratio between the Zeeman splitting and the Doppler width. (3) As the temperature-density stratification inside the fluxtubes is greatly different from that of the external atmosphere, it is not clear to which geometrical height the field-strength determinations refer. Since the field strength decreases with height, the derived field strengths are intimately linked to the opacity structure of the fluxtubes.

We will come back to the issue of the opacity structure in Sect. 3.4 and 5.2, and focus the discussion here on the first problem, the field-strength distribution. When the line-ratio technique was first applied to derive the intrinsic field strengths, there did not exist any observational constraints on the field-strength distribution, so the interpretation was made with different assumptions for the horizontal field-strength variation across the fluxtube at the atmospheric level where the line wing is formed in the fluxtube (Stenflo, 1973). For a rotationally symmetric element, the more peaked the magnetic cross-section is, the higher is the field amplitude needed to reproduce the slope in the line-ratio scatter plot. Thus, in a quiet region at disk center, the field strength at the fluxtube axis was found to be 2.3 kG assuming a gaussian cross section, while it was 1.1 kG for a rectangular cross section (Stenflo, 1973). The line-ratio data alone do not allow us to distinguish between these cases.

With the recent possibility of recording fully spectrally resolved polarized line profiles with the FTS in the infrared, observational constraints on the field-strength distribution are now available, and they favour the rectangular, horizontal cross section. A field-strength distribution gives rise to Zeeman line broadening, which becomes prominent in the infrared due to the wavelength dependence of the Zeeman splitting. The problem is however that the non-thermal Doppler broadening inside the fluxtubes is significantly larger than in the external atmosphere, and contributes to the broadening of the Stokes V profiles. The trick to separate the effects of Doppler and Zeeman broadening from each other is to use a kind of "infrared line ratio", with two lines having greatly different Zeeman sensitivities. This is illustrated in Fig. 7, from Zayer et al. (1989). The Stokes V profiles (solid curves) in a network element at disk center for the two Fe I lines at 15 648.54 Å (Landé factor 3.0) and 15 822.81 Å (Landé factor 0.75) are displayed and compared with the corresponding profiles predicted with the assumption of a single-valued field (dashed curves). The two lines have practically the same excitation potential of the lower atomic level (5.43 and 5.64 eV, respectively) and similar line strengths, and are therefore formed at

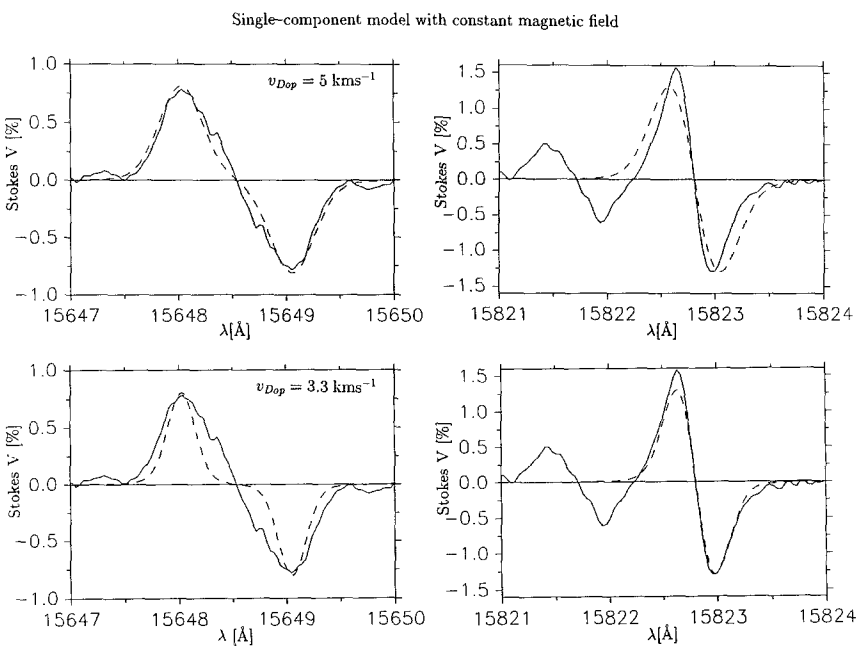

Fig. 7. Illustration of how models without magnetic broadening (distribution of field strengths) are unable to account for the observed Stokes V profiles in the infrared. When as in the top diagrams the Doppler broadening is adjusted to fit the "Zeeman" line (to the left), the model profile (dashed) is too broad in the case of the "Doppler" line (to the right). When instead the Doppler broadening parameter is adjusted to fit the "Doppler" line, the model profile for the "Zeeman" line becomes much too narrow. It turns out that the required Zeeman broadening can be entirely accounted for by the height gradient of the magnetic field, without any need for a horizontal variation of the field strength. From Zayer et al. (1989)

almost the same height in the atmosphere. Due to the small Landé factor of the second line (the "Doppler line"), Zeeman broadening is insignificant as compared with the Doppler broadening, whereas the first line (the "Zeeman line") is quite strongly Zeeman broadened.

Fig. 7 shows that when no field-strength distribution is introduced, and the Doppler broadening is adjusted to fit the width of the "Zeeman" line, then the predicted width of the "Doppler" line is too large (upper portion of Fig. 7). When on the other hand the Doppler broadening is adjusted to fit the Doppler line, then the predicted width of the Zeeman line is much too small. This clearly demonstrates that Zeeman broadening due to a field-strength distribution must be invoked to explain the observed profiles.

The field-strength distribution can be both across the horizontal plane as well as in the vertical (along the line of sight) direction (due to a height gradient of the field). The height gradient can be constrained by demanding that the observed 5247–5250 line ratio, which refers to a higher atmospheric level than the infrared lines, should be satisfied simultaneously, whereby a magnetohydrostatic fluxtube in the "thin tube approximation" (which is valid over this height range) can be used for a self-consistent model interpretation. Using this approach,

Zayer et al. (1989) found that the height gradient alone gave rise to the right amount of Zeeman broadening, without the need for any horizontal field-strength distribution (which would tend to make the Zeeman broadening too large).

This result suggests both that the field strength at lower photospheric levels is almost constant horizontally across the fluxtube, and that the field strength in different fluxtubes within the spatial resolution element (which generally does not isolate a single flux element, but a cluster of elements) is practically identical in the different fluxtubes. Since the observations used by Zayer et al. (1989, from Stenflo et al., 1987 b) were obtained in a fairly strong plage, the resolution element (circular, with a diameter of 5 arcsec) certainly included a large number of fluxtubes. These conclusions support the validity of the two-component model approach.

An indication that there nevertheless may be some small horizontal variation of the magnetic field comes from the two "notches" occurring in the Zeeman-line Stokes V profile between the V maximum and the central zero crossing, symmetrically on both sides of the line center in Fig. 7. These notches also show up at a recording closer to the solar limb, at $\mu = \cos\vartheta = 0.61$. The model fitting efforts suggest an explanation in terms of a small opposite-polarity return flux of intermediate strength, returning about 3–7 % of the flux to the immediate surroundings of the fluxtube (Zayer et al., 1989). A similar picture was arrived at in line-ratio analysis by Frazier and Stenflo (1972).

The above discussion has provided some support for the validity of the two-component model, and for the idea that the fluxtube properties are "unique" in the sense that they do not vary much from one fluxtube to the other. As however the statistics with the FTS approach is very limited, with only few spatial points on the solar disk being sampled, the main evidence for the uniqueness of the fluxtubes came already with the magnetograph line-ratio scans (Howard and Stenflo, 1972; Frazier and Stenflo, 1972), since thousands of spatial points over the whole range of filling factors could be sampled (at the expense of spectral information — only two wavelengths were used). The arguments used can be illustrated with reference to the scatter plot in Fig. 5.

First we note that the scatter of the points around the straight line in Fig. 5 does not exceed that due to random instrumental noise, and, within the limits of error, the regression curve appears to be a straight line with the same slope, independent of apparent field strength (filling factor). The size of the scanning aperture used in this as well as in a previous investigation was $2.4 \times 2.4\,\text{arcsec}^2$. If there were a distribution of fluxtube properties, with weak-field and strong-field fluxtubes occurring in different regions on the solar disk, we would expect a continuous and broad population of the scatter-plot diagram, with points from the weak-field regions falling along the 45° line, those from kG regions along a line similar to that of Fig. 5, while points from regions with intermediately strong fields would be spread over the area between these two lines. The circumstance that all points appear to follow the kG line indicates that there are no significant weak-field fluxes.

These conclusions have been verified using line-ratio observations with higher signal-to-noise ratio to more clearly discriminate between weak and strong fields

also in the case of small filling factors (Stenflo and Harvey, 1985). Although the statistical sample was limited since full Stokes V line profiles were recorded, all determined line ratios (on a scale where 1.0 represents the weak-field case) were found to be smaller than 0.82 (with error bars ≈ 0.02). If there had been weak fields, line ratios of unity would have been found. No points fell in the interval 0.82 – 1.0. When the filling factor was decreased by a factor of 6, the intrinsic field strength was found to decrease by only about 20 % .

It could be argued that most points on the solar disk have filling factors too small for the line ratio to be determined, and that the fields there could be intrinsically weak. Since these points (which cluster around the origin in a scatter-plot diagram like Fig. 5) together represent a large fraction of the solar surface, they might also carry a significant fraction of the total magnetic flux through the surface. The conjecture that these points generally represent weak fields can however be tested and disproved by summing the absolute values of the fluxes measured at all the disk positions for each of the two lines (5247 and 5250), and forming the ratio of the total fluxes over the whole scanned area. If a substantial portion of the total flux were intrinsically weak, the B_{5250}/B_{5247} ratio for the total fluxes should be closer to unity than the ratio obtained from the slope in the scatter-plot diagram (which is more determined by areas with larger filling factors). Within the error limits, the slope and the total flux ratio are found to be the same, suggesting that all the flux is in strong-field form. The error bars would permit some intrinsically weak fields to exist, but these fields could account for at most 10 % of the total flux. More than 90 % of the flux that is recorded by magnetographs with 2.4 arcsec resolution is therefore carried by strong (kG) fields with small magnetic filling factors (Howard and Stenflo, 1972; Frazier and Stenflo, 1972; Stenflo, 1976). In Sect. 4 we will further consider the physical meaning of this remarkable conclusion.

3.4. Fluxtube temperature structure

Before the advent of Stokes V fluxtube diagnostics, the temperature stratification in magnetic regions was represented by "facular" models, constructed on the basis of observations of unpolarized radiation. All such models are however dependent on the spatial resolution used, since even with the best available resolution the magnetic filling factor is small, which means that the spectral information is greatly contaminated by the radiation from the atmosphere outside the fluxtubes. As mentioned above, it is only by using polarized radiation that this "stray light" problem can be avoided. The Stokes V spectrum, e.g. that of Fig. 6, comes exclusively from the fluxtubes, regardless of how small the filling factor is. The Stokes I spectrum on the other hand has contributions from the whole spatial resolution element. If the filling factor is small, Stokes I is mainly representative of the surrounding, "non-magnetic" atmosphere.

Temperature weakening of Stokes I line profiles were found in faculae by McMath et al. (1956), Sheeley (1967), and Chapman and Sheeley (1968). As they originate in the small filling-factor fraction of the area, their effect on Stokes I

is found to be proportional to the amount of flux measured (i.e., to the filling factor). Only in Stokes V can the effect be observed by its full amount.

Fig. 6 illustrates the effect of temperature weakening on Stokes V. The line Fe I 5250.65 Å has a Stokes V amplitude that is even larger than its neighbouring 5250.22 Å line, although its Landé factor is only half as large (1.5). Analysis shows that about half of this effect is due to Zeeman saturation in the 5250.22 Å line, while the other half is due to differential temperature weakening. The temperature line weakening increases with decreasing excitation potential of the lower atomic level. The "magnetic line pair" 5247.06 and 5250.22 Å has low excitation potentials, 0.09 and 0.12 eV, respectively, and are therefore quite temperature sensitive. As this sensitivity is practically the same for both lines, it will cancel out when forming the Stokes V line ratio, allowing the differential effect of Zeeman saturation to be isolated. The 5250.65 Å line on the other hand has an excitation potential of 2.20 eV, and is thus significantly less temperature sensitive.

As the Zeeman saturation in both the 5247.06 and 5250.65 Å lines is quite small, the Stokes V ratio between these two lines is dominated by temperature effects. This line pair is therefore suitable for use as a "thermal line ratio" (Stenflo et al., 1987 a).

The first use of Stokes V line weakening to derive the temperature stratification inside the fluxtubes independent of the spatial resolution of the observations was made by Stenflo (1975), based on magnetograph line-ratio data. With the advent of the FTS polarimeter, a much more powerful approach was developed, using many hundreds of simultaneously recorded Stokes V line profiles as collective constraints on the possible fluxtube temperature stratifications (Solanki and Stenflo, 1984, 1985). Such a statistical approach had previously been developed by Stenflo and Lindegren (1977) with the objective of searching for tiny magnetic line broadening effects due to a possible turbulent magnetic field. The method consists of first characterizing the line profiles by some suitable small set of parameters (line width, depth, etc.), extracting these parameters from the hundreds of selected lines in the observed spectrum, and then determining how these empirical parameters depend on line strength, excitation potential, Landé factor, and wavelength by making fits with a relatively sophisticated regression equation (the proper form of which is first explored by making various scatter plots of the line parameters vs. line strength, excitation potential, etc.). This powerful diagnostic approach is now being applied to other stars as well (Mathys and Stenflo, 1986; Mathys and Solanki, 1989).

The proper choice of parameterization for the Stokes V profiles is greatly facilitated by the circumstance that Eq. (4), retaining only the first term in the Taylor series expansion, is a good approximation for most lines (since very few lines in the visible, like Fe I 5250.22 Å, are significantly affected by Zeeman saturation). If we integrate the Stokes V profiles with respect to wavelength, $\int^{\lambda} V(\lambda')\,\mathrm{d}\lambda'$, and divide by $\Delta\lambda_H$, we obtain a quantity, designated $I_V(\lambda)$ by Solanki and Stenflo (1984), which is a good approximation of I_α. We have thus practically retrieved the Stokes I profile of the magnetic elements without spatially resolving them. The I_V profiles can then be parameterized in the same

way as any unpolarized I profiles. This means that the parameterization scheme of Stenflo and Lindegren (1977) can be applied to both Stokes I and V spectra in a unified way.

The Stokes V profiles of about 400 Fe I and 50 Fe II lines have been used to simultaneously constrain the possible temperature stratification inside the fluxtubes (Solanki and Stenflo, 1984, 1985). The interpretation has so far been restricted to 1-D (using one line of sight) LTE models, using the so called thin tube approximation to couple the magnetic field in a self-consistent way to the density-temperature structure. As the spectral lines used are distributed over all line strenghts, they span a range of heights of formation throughout the photosphere to the temperature minimum. The derived temperature difference $\Delta T(\tau_{5000})$ between fluxtube interior and exterior *at equal optical depth* shows a fluxtube temperature excess at all optical depths by amounts of 400–800 K. This does not imply that the temperature is higher at equal *geometrical* heights — on the contrary the fluxtubes are rather expected to be somewhat cooler at lower levels due to partial inhibition of the convective energy transport. It is further found that the fluxtubes in the quiet network are hotter (i.e., $\Delta T(\tau_{5000})$ is larger) by a few hundred degrees in the lower photospheric layers as compared with fluxtubes in plages (Solanki and Stenflo, 1984, 1985; Solanki, 1987 a, b). This can be understood in terms of stronger inhibition of convection by larger flux concentrations (Schüssler, 1987; Knölker and Schüssler, 1988).

Since the derived temperature structure is based on a number of quite restrictive assumptions, and the uniqueness of these results has not been sufficiently explored, the empirical ΔT curves should be regarded as tentative. The interpretation lags far behind the observations. Thus the extensive FTS Stokes V data set with the center-to-limb variations (Stenflo et al., 1987 a; Pantellini et al., 1988) has not been sufficiently exploited yet (which requires integration over many different lines of sight, i.e., "1.5-D" radiative transfer). This is also the case for the infrared (around 1.6μm) FTS data (Stenflo et al., 1987 b), which extend the height coverage to somewhat deeper levels. Fully self-consistent magnetohydrostatic models, like those of Steiner et al. (1986) and Steiner and Pizzo (1989), need to be coupled to the diagnostic problem. Non-LTE effects in fluxtubes, like those explored by Solanki and Steenbock (1988) and Stenholm and Stenflo (1978), need to be taken into account.

One approach of considerable future potential is the application of inversion techniques to derive the vector magnetic field (Skumanich and Lites, 1987; Lites et al., 1987, 1988) or the temperature structure (Keller, 1988) from Stokes data. For instance, in the work of Keller (1988), the height variation of the temperature is characterized by a small number of free parameters to be determined by an iterative least-squares fitting procedure. Each parameter set defines a self-consistent magnetohydrostatic model, which can be used to compute any set of chosen Stokes observables. These observables have to be selected in a careful way, mainly as various Stokes V line ratios between high and low excitation potential lines, Fe I and Fe II lines, weak and strong lines, etc., to respond as "orthogonally" as possible to variations of the free model parameters. Minimizing the χ^2 between

the observed and synthetic (computed) observables leads to a determination of the model parameters, including their error bars.

Finally another approach with fundamental implications for our understanding of the heating processes in the fluxtubes should be mentioned. Ayres et al. (1986) found that their FTS recordings (without polarization optics) made simultaneously in the infrared molecular bands of carbon monoxide at 2.3 and 4.7 μm, and in the ultraviolet Ca II K line at 3933 Å, could not be interpreted with any single model atmosphere, but that a two-component model (hot and cold) at the level around or somewhat above the temperature minimum was needed, indicating what they called a "thermal bifurcation" of the solar atmosphere. The most natural scenario for explaining this result is that the chromospheric temperature rise only takes place inside the magnetic regions. In the external atmosphere the temperature decreases monotonically with height, until the overlaying magnetic canopy is reached, where a transition to the hot magnetic region occurs.

3.5. Fluxtube dynamics

As was the case for the temperature structure, we also need to use polarized (Stokes V) spectral data to avoid contamination of the spatially unresolved fluxtubes by radiation from the surrounding atmosphere when studying the fluxtube dynamics, since the dynamical state is expected to differ greatly between the fluxtube interior and exterior. The most powerful approach is again the statistical one, determining regression relations for 400 Fe I and 50 Fe II lines (Solanki, 1986). The main diagnostic parameters for the fluxtube dynamics are:

- The Stokes V amplitude and area asymmetries.
- The Stokes V zero-crossing wavelength.
- The difference in width between the I_V and I line profiles.
- The center-to-limb variations of the above parameters.

The Stokes V asymmetry means that the blue and red wing portions of the profiles are not simple mirror images of each other. It can be seen in Fig. 6 that the V amplitude in the blue wing is larger than that in the red wing. This is the case for all spectral lines, all over the solar disk, regardless of the magnetic polarity. The *area* of the blue-wing portion of V is also larger than that in the red wing at disk center, but this area asymmetry changes sign near the solar limb (Stenflo et al., 1987 a). Such amplitude and area asymmetries in Stokes V can arise only if there are correlations between the magnetic and velocity field gradients within the spatial resolution element. For an area asymmetry to result, such correlations have to occur along the line of sight, as first noted by Illing et al. (1975). Both the amplitude and area asymmetries have their maximum values for intermediately strong lines, and vanish when the line strength goes to zero.

If we consider the Stokes V asymmetry by itself, disregarding the other observables constraining the dynamics, it can be explained in terms of downdrafts inside the fluxtubes, the downdraft velocity increasing with height as required by mass conservation, at the same time as the magnetic field strength decreases with height, due to the decreasing external gas pressure (Solanki and Pahlke, 1988).

However, any such model with a unidirectional mass flow inside the fluxtubes must be discarded, since all of them predict a net Doppler shift, in the case of the downdraft model a redshift, of the Stokes V zero-crossing wavelength, of the order of one km s^{-1} or more. Observations show that the actual shift of the zero crossing is smaller than about 0.25 km s^{-1} (Stenflo and Harvey, 1985; Solanki, 1986), which effectively rules out all models based on quasi-stationary mass flows *inside* the fluxtubes.

Recently, however, Grossmann-Doerth et al. (1988), using ideas first introduced by van Ballegooijen (1985), have shown that downdrafts immediately *outside* the fluxtubes are capable of producing the observed type of V asymmetry without any zero-crossing wavelength shift. It may appear paradoxical that the non-magnetic external atmosphere can have an influence on the polarization that originates inside the fluxtube, but this is the case when the line of sight crosses the fluxtube boundary. As the fluxtube cross section diverges with height, a "canopy" overlying the field-free region develops. By comparing the appearance of features in magnetograms recorded in photospheric and chromospheric spectral lines, Giovanelli (1980) and Jones and Giovanelli (1983) concluded that these canopies become quite extensive as low down as the temperature minimum.

If the unpolarized radiation that enters the fluxtube from below across the canopy region is Doppler shifted due to downdrafts in the non-magnetic atmosphere, it will affect the circular polarization being formed in the non-moving magnetic region above, such that the blue-wing polarization is enhanced relative to the red-wing polarization, leading to the observed Stokes V asymmetry without Doppler-shifting the V profile. This picture of surrounding downdrafts is quite satisfying, since it has long been known that there is a close spatial correlation between magnetic flux concentrations and downdrafts observed via the Doppler shifts of the I profiles (Simon and Leighton, 1964; Tanenbaum et al., 1969; Frazier, 1970). Quasi-stationary downdrafts inside the fluxtubes would be difficult to understand from a theoretical point of view, since mass conservation would lead to a drainage of the corona in a matter of minutes (diffusion across the field lines from the denser layers in the external atmosphere is too slow to provide the needed mass flux).

Although no systematic, unidirectional flows seem to exist inside the fluxtubes, large-amplitude non-stationary motions must take place there, as evidenced by the observed non-thermal or macroturbulent Doppler broadening of the I_V profiles (wavelength-integrated V profiles, cf. Sect. 3.4). The macroturbulent velocities in the fluxtubes increase with line strength (height in the atmosphere), from 1.8 km s^{-1} for weak lines to about 3.5 km s^{-1} for strong Fe I lines (Pantellini et al., 1988). These velocities are practically independent of center-to-limb distance, i.e., they are as large in the direction perpendicular to the field lines as parallel to them. The most natural candidate process to explain these mass motions are fluxtube waves, but they need to have quite large amplitudes. The dissipation of these waves may be an important source of heating, not only for the fluxtubes, but also globally for the corona, since the corona is filled with magnetic fields that lower down converge to form the small-scale photospheric fluxtubes.

3.6. Fluxtube sizes

The size distribution of the flux elements has been discussed in Sect. 2.3 from a general point of view. It was argued that larger flux elements emerging at the surface from the solar interior become shredded due to the interchange instability and the action of the turbulent motions, and that the flux thereby is transferred rather quickly to the small-scale end of the size spectrum. This process must be at least partially stopped by some counteracting concentrating or stabilizing mechanism before sizes at which ohmic dissipation becomes dominant are reached, otherwise we could hardly account for the observation that more than 90 % of the overall flux at any given time is in the form of strong (kG) fields. The scale at which further fragmentation is balanced or stopped should also represent the typical size scale of the magnetic elements outside sunspots.

Two main questions need to be answered by observations: (1) Is there really a preferred size scale, or is there a wide range of various sizes ? (2) If there is a preferred size, how large is it ? These questions cannot be resolved by direct observations alone, as has been tried in the past using bright points as proxies for the magnetic field (Spruit and Zwaan, 1981), since the observed elements may actually represent unresolved clusters of smaller elements. Also the derived magnetic filling factors do not give us much guidance (except for providing upper limits to the sizes), since a given filling factor may be due to the accumulated effect of many elements.

These problems may at a first glance seem insurmountable, but they may in fact be solved by indirect, spectral methods, independent of the spatial resolution used in the observations. Zayer et al. (1989) have recently done just that, and provided tentative answers to the above two questions. Their answers are: (1) Yes, there appears to be a characteristic or "unique" size scale. (2) The value of the derived size of the cross section at the $\tau_{5000} = 1$ level depends critically on the assumed geometry of the fluxtubes. For a rotationally symmetric "funnel" geometry of a solitary fluxtube (without neighbours), the fluxtube diameter is found to be about 300 km. If instead a slab geometry is assumed, the slab thickness is found to be 60–100 km. The reason for this dependence on assumed geometry will be explained below. Although these results are very tentative and should be used with great caution, the work of Zayer et al. (1989) indicates how more definite answers to these questions may be obtained in the future.

The spectral information used for the size determination was the same as in the case of the horizontal and vertical magnetic-field distributions described in Sect. 3.3, except that the emphasis was on analysis of data obtained far from the center of the disk. Thus the FTS Stokes V profiles of the infrared lines Fe I 15,648.54 and 15,822.81 Å were used, in combination with the line ratio pair Fe I 5247.06 and 5250.22 Å. Since the fluxtubes diverge with height and develop "canopies", the radiative-transfer interpretation requires the use of many lines of sight at different distances from the fluxtube axis, even when only vertical fluxtubes at disk center are considered. At disk center, however, the size of the fluxtube cross section does not affect the relative shapes of the Stokes V line profiles.

The situation is different for fluxtubes away from the center of the disk, however. Since they are seen at an angle, their optical thickness in the horizontal direction becomes important. This optical thickness determines how much of the line-forming region is included inside the fluxtube, and how much falls outside it. The optical thickness scales with the fluxtube diameter, leading to a diameter-dependence of the observed V profiles. The requirement that the above-mentioned model should simultaneously satisfy both the disk center ($\mu = 1.0$) and $\mu = 0.61$ observations provided sufficient constraints to derive the diameter values mentioned above.

The reason why it makes such a difference whether a slab or a funnel model is assumed has to do with the rate of divergence with height of the fluxtube cross section needed to satisfy flux conservation. As slabs only diverge in one dimension, their linear cross section scales with the field strength B as B^{-1}. The funnels on the other hand, diverging in two dimensions, have cross sections scaling as $B^{-1/2}$. Due to the considerably more rapid divergence of the slabs, a substantially larger portion of the inclined lines of sight reside inside them as compared with the funnels. This picture might change if we consider a cluster of fluxtubes instead of the solitary fluxtubes in the treatment of Zayer et al. (1989), since a single line of sight may pass through more than one fluxtube. As the height divergence critically depends on the internal temperature and density structure, a more definite determination of the fluxtube sizes is quite a complex problem, although within our reach.

Once the geometry of the fluxtubes has been fixed, it can be concluded that all the fluxtubes inside the spatial resolution element must have similar sizes. If there were a substantial spread in the size distribution, this would lead to additional Zeeman broadening that cannot be accounted for by adjusting the other model parameters (cf. Sect. 3.3). As the observational data had been obtained in a strong plage with fairly large filling factor, the results of Zayer et al. (1989) indicate that larger filling factors do not imply larger fluxtube sizes, but rather a larger number density of similar fluxtubes. Such a picture is supported by the quite small filling-factor dependence of the magnetic and thermodynamic properties derived by the line ratio or statistical multi-line techniques. Knölker and Schüssler (1988) also reach the same conclusion (using 2-D MHD simulations) that larger filling factors correspond to clusters of smaller magnetic elements. They further conclude that the joint effect of the fluxtubes is to reduce the convective energy transport also between the fluxtubes, making larger clusters appear dark in continuum radiation observed with moderate spatial resolution.

4. "Hidden" or "turbulent" magnetic flux

Our previous conclusion that more than 90 % of the flux through the solar photosphere occurs in strong, kG form implies, when compared with the average flux densities observed, that the global average of the magnetic filling factor is only a fraction (a few tenths) of one percent. The question then naturally arises what goes on in the remaining 99 % of the photosphere. For physical reasons

this, what is sometimes called the "non-magnetic" region, cannot be entirely field free, but fields must occur everywhere at some level of the field strength. The main questions we need to address are therefore: (1) What is the characteristic strength of this "invisible" field ? (2) How is it structured ? (3) How is it related to the concentrated kG fields ?

A large amount of magnetic flux could be "hidden" from view in magnetograms if the opposite-polarity fluxes are mixed on a scale smaller than the spatial resolution element, such that the integrated net flux over the resolution element is zero. If this is the case, the resulting Zeeman-effect circular polarization also averages out to zero, both in magnetograms and FTS spectra.

Although this type of mixed or "turbulent" magnetic flux escapes detection by the longitudinal Zeeman effect, it can reveal itself in other ways. The following methods have been used to constrain the properties of the "hidden" magnetic field:

- Zeeman broadening of unpolarized line profiles.
- Hanle-effect depolarization.
- Transverse Zeeman effect.
- Direct observations with the highest possible spatial resolution.

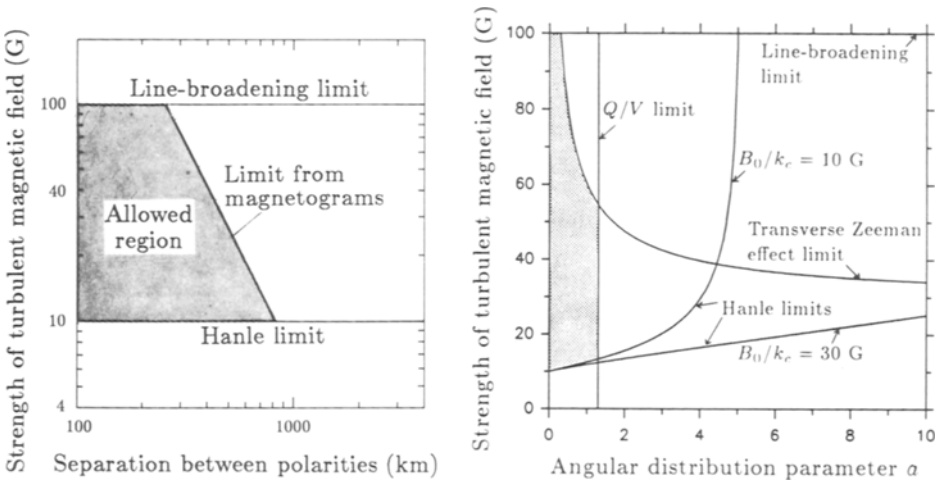

Fig. 8. Observational limits on the "turbulent" magnetic field or "hidden" magnetic flux on the sun. In the diagram to the left, the strength of the turbulent field is given vs. the characteristic separation scale between the opposite magnetic polarities (Stenflo, 1984), in the diagram to the right vs. a parameter a characterizing the angular distribution of field vectors (Stenflo, 1987). $a = 0$ corresponds to an isotropic distribution, which becomes more and more peaked around the vertical direction with increasing value of a. The observationally allowed regions are indicated by the shaded areas in the two diagrams (see text)

The results that have been obtained so far by applying these methods are summarized in Fig. 8, from Stenflo (1984, 1987). The "turbulent" field has not yet been positively identified, but it has been possible to substantially constrain its properties. The observationally allowed regions are indicated by the shaded areas

in the two diagrams of Fig. 8. The left diagram gives the strength of the field as a function of the typical spatial separation between the magnetic polarities in the field pattern. If the separation is large enough, the turbulent field will be revealed directly in magnetograms if this distance can be spatially resolved. The seeing-limited spatial resolution depends on the integration time used to obtain a certain polarimetric accuracy, and the polarimetric accuracy is improved upon by making trade-offs with the size of the instrumentally selected spatial resolution element, to enhance the photon statistics. Therefore the magnetogram resolution limit, based on results of Harvey (1977) and Tarbell et al. (1979), is slanted in Fig. 8. The allowed region does not end at the left edge of the diagram (separation 100 km) but continues unconstrained indefinitely towards smaller sizes (a theoretical lower limit is likely to occur at the much smaller ohmic diffusion scale).

The Zeeman line-broadening limit is based on a statistical analysis of 400 unblended Fe I lines, for which the spectral line width is represented by a regression equation as a function of line strength, excitation potential, wavelength, and Landé factor (Stenflo and Lindegren, 1977). The regression coefficient of the term that describes the Landé factor dependence contains the information on the strength of the turbulent field. The one-sigma error bar for the determined regression coefficient translates to the upper limit of 100 G for the turbulent field. A 3-sigma limit would correspond to a field strength $\sqrt{3}$ times as large.

The methods applied in the right diagram of Fig. 8 are resolution independent, and no length scale is involved. The angular distribution parameter on the abscissa characterizes the statistical orientation of the field vectors of the turbulent field. The field direction can be described in terms of the inclination angle ϑ_B with the vertical direction in the atmosphere, and by the azimuthal angle (in the horizontal plane). Making the natural assumption that the azimuthal angles are randomly distributed, the full vector distribution can be described in terms of the distribution of inclination angles. For the purpose of deriving crude constraints, we let this distribution be expressed as being proportional to $\cos^a \vartheta_B$, thus being characterized by the single free parameter a, the angular distribution parameter of Fig. 8. $a = 0$ corresponds to an isotropic distribution of field vectors. As the value of a is increased, the distribution becomes more and more peaked around the vertical direction (which may happen if buoyancy forces dominate).

The Hanle limit, which is a *lower* limit, is based on a very tentative analysis of the depolarization, caused by a turbulent magnetic field, of the linear polarization formed by coherent scattering in the core of strong spectral lines (Stenflo, 1982). The scattering polarization, which increases when approaching the solar limb, has been recorded throughout the solar spectrum (3165–9950 Å) (Stenflo et al., 1980, 1983 a, b). By comparing the observed core polarization with what would be expected in the absence of magnetic fields, the strength of the depolarizing magnetic field can in principle be found. The problem is that the calculation of the expected core polarization is a very complex undertaking, since it depends on the geometry of the radiation field throughout the atmosphere, the collisional depolarizing rate, and the detailed non-LTE transfer of polarized, scattered radiation.

A full treatment of this problem has not yet been done, but by considering the physics of the last scattering process, crude estimates of the magnetic and collisional depolarizations have been made for a number of polarized lines, leading to the *lower* limit to the strength of the turbulent field of about 10 G in the case of an isotropic distribution of field vectors (Stenflo, 1982), as illustrated in Fig. 8. When other field distributions are assumed, i.e., other values of the angular distribution parameter a, the Hanle depolarization limit changes in a way that depends on the parameters of the atomic transition used, or more specifically, on B_0/k_c, where B_0 is the field strength for which the Larmor precession period equals the natural life-time of the upper level, and k_c is the depolarizing collision rate relative to the total rate (collisions plus radiative transitions). B_0/k_c is thus the field strength for which the Larmor precession period equals the *effective*, undisturbed (by collisions) life-time of the upper level. The Hanle results should be used with great caution, and have been entered here more to indicate the important diagnostic potential of the method, although it is difficult to apply in practice.

Finally we have the limits in the right diagram of Fig. 8 based on the use of the transverse Zeeman effect, both by itself, as well as in ratio with the longitudinal Zeeman effect (Q/V ratio in Fig. 8). The main principle that is used to obtain these limits is a symmetry property of the transverse Zeeman effect: When the field direction is reversed, the circular polarization (longitudinal Zeeman effect) changes sign, whereas the linear polarization (transverse Zeeman effect) is invariant (a sign change occurs for 90° rotation of the transverse vector component, but a 180° rotation brings back the original value). A distribution of field vectors with the field strength invariant with respect to rotations around the vertical direction, but with zero net flux through the solar surface, may produce, due to the above symmetry property, a net, non-zero Stokes Q tranverse Zeeman-effect signal after integration over all the field directions. These considerations applied to Stokes Q and V line profile recordings with high signal-to-noise ratio led to the limits of Fig. 8 (Stenflo, 1987).

In summary, the properties of the "turbulent" field have tentatively been constrained as follows:

- The field strength is likely to be in the range 10–100 G.
- The distribution of field vectors is almost isotropic.
- The characteristic scale size is smaller than one arcsec.

The nature of this "turbulent" background field and its relation to the concentrated kG fields are not known, but a possible scenario relating the mixed-polarity background field to "$\cup$-loop" field configurations has been presented by Spruit et al. (1987). A $\cup$-loop occurs when two kG regions of opposite-polarity magnetic flux in the photosphere are connected below the solar surface through a largely horizontal subsurface loop. Sections of this buoyant loop are brought to the surface through the action of local turbulent eddies, thereby generating a mixed-polarity background field between the vertical kG fluxtubes. Reconnection in the upper part of the emerged sections changes the topology from that of a connected "sea serpent" to that of a number of separated, smaller closed loops,

which decay when the "wrapping up" of the field by turbulent turnover reduces the length scale to values at which ohmic diffusion becomes important (of the order of 1 km). The net effect is removal of flux from the sun, at a rate that globally might account for the observed solar-cycle variations.

5. Origin of the small-scale fields

5.1. Convective collapse

Weak, vertical magnetic flux in the solar surface layers may be spontaneously concentrated into a kG state by the instability mechanism of convective collapse (Parker, 1978; Spruit, 1979; Spruit and Zweibel, 1979; Unno and Ando, 1979). The principle is simple. The interior of a magnetic flux region is partially thermally insulated from its surroundings due to inhibition of convective energy transport perpendicular to the field lines, and to the reduction of the lateral radiative heat exchange caused by the finite optical thickness of the region in the horizontal direction. If we introduce a small disturbance in the form of a weak downdraft, then, for a thermally insulated flow, an adiabatic temperature gradient will be set up. The temperature gradient of the external medium in the upper convection zone is however superadiabatic, i.e., steeper. The disturbance thus leads to a temperature deficit, accompanied by a pressure deficit, in the flux region. Pressure balance is restored by contraction of the flux region, which amplifies the downdraft disturbance, leading to further contraction, etc. The result is a spontaneous collapse of the magnetic region until a new equilibrium state can be reached. This new equilibrium is a strong-field state for which the magnetic pressure is comparable to the ambient gas pressure, so that the magnetic field can resist further compression. As the gas pressure in the photosphere corresponds to the magnetic pressure of a field of strength 1–2 kG (Stenflo, 1975), it is natural to expect that the collapsed end state will have field strengths in that range.

A linear stability analysis by Spruit and Zweibel (1979) has indicated that flux regions with $\beta = 8\pi P/B^2 \lesssim 1.8$ are stable against convective collapse, while those with smaller β are unstable. This critical value of β corresponds to a field strength of about 1.3 kG in the photosphere, in good agreement with the observed field strengths. Non-linear calculations (Hasan, 1984, 1985; Venkatakrishnan, 1985, 1986) confirm this picture but indicate that the final state is not a static one. Instead, lateral radiative exchange with the surroundings leads to overstability, with stationary large-amplitude oscillations having computed periods of the order of 1000 s. Such oscillations might contribute to the observed non-thermal broadening, possibly also to the observed asymmetry, of the Stokes V line profiles (cf. Sect. 3.5).

In spite of the great attraction of the convective collapse scenario, this process has never been observed on the sun, although its final state is identified with the kG flux elements we see in the photosphere. One reason for this is that the collapse process is fast, of the order of 5–15 min, and the combined time and spatial resolution of existing instruments is simply not sufficient for this process to be observed. As the flux is expected to remain in the strong-field state for a

time that is long in comparison with the collapse time, we would expect to see most of the flux at any given time in the strong-field state, which we do (cf. Sect. 3.3). In addition, it is likely that the flux is already in a preconcentrated state when it emerges from the solar interior to the surface, and that the process of convective collapse thus only has to take care of the final concentration, from e.g. 0.1 to 1–2 kG. The preconcentration is achieved by the well-known mechanism of convective expulsion of flux (Weiss, 1966; Peckover and Weiss, 1972).

In convective expulsion, the frozen-in magnetic field is carried to the boundaries of the convective or turbulent cells. This field concentration to the cell boundaries should occur throughout the convection zone of the sun. In the photosphere, however, due to the low gas pressure there, the dynamic pressure of the turbulent flow is insufficient to produce kG fields, and can only amplify fields to the 100 G range. Therefore convective collapse is needed to explain the kG fields. Deeper inside the convection zone the situation is however entirely different. There convective expulsion is capable of generating fields much stronger than the observed surface fields. When these fields rise to the surface they expand and therefore weaken due to the rapidly decreasing gas pressure in the surrounding medium, but we may still expect the fields to be in a fairly fragmented, intermittent form when they emerge.

Extensive 3-D numerical simulations of the solar granulation, solving the time-dependent MHD equations in the stratified solar atmosphere, with radiative exchange and magnetic fields, have been carried out by Nordlund (1983, 1986). These simulations, which are very successful in reproducing the observed properties of the solar granulation, also show the evolution of the magnetic flux, including both the processes of convective expulsion and in principle also convective collapse. Due to computer limitations, however, the spatial grid size is not sufficient for resolving the final, collapsed state of the flux. The kG flux elements therefore have to be treated separately by more idealized models (basically reducing the dimensionality of the problem from 3 to 2), as will be described in the next section.

Nordlund's simulations also show how the magnetic flux patches are neutral to shape deformations, fragmentation, and coalescence. This leads us to some comments on the stability of the field structures. Meyer et al. (1977) showed that while small fluxtubes are unstable against fragmentation by the interchange or flute instability, larger fluxtubes are stabilized by buoyancy. Schüssler (1984) pointed out that small fluxtubes will be stabilized against fragmentation if they are surrounded by a vortex flow (similar to a tornado) with velocities of the order of 1 km s^{-1} or larger. Due to the "bathtub effect", such a flow naturally develops in the downdraft regions at the convective cell boundaries, where the flux is concentrated (cf. Stenflo, 1975). Recently a vortex flow of this kind has been observed directly in a high-resolution movie of the solar granulation (Brandt et al., 1988).

It should finally be recalled that the process of convective collapse gets suppressed when the lateral radiative exchange due to small horizontal optical thickness becomes efficient, since then the interior of the flux region is no longer thermally insulated from its surroundings. This is expected to occur at

scales of the order of 100 km in the photosphere. At these scales we expect a complicated interaction and balance between a number of different physical processes: Fragmentation and coalescence, vorticity development and decay, radiative exchange and wave heating (leading to flux dispersal), recollapse, ohmic decay of current sheets, etc. Another consequence of these processes is that two field lines that are part of the same flux fragment at a given time may end up in widely separated flux fragments later. The field line connections below or above will as a result become increasingly tangled with a number of field line crossings (topological "catastrophes"), at which reconnection can take place to simplify the field configuration.

5.2. Modelling the equilibrium state of the concentrated fields

The first comprehensive model of the equilibrium state of small-scale photospheric fluxtubes is due to Spruit (1976), who used the thin-tube approximation in rotationally symmetric geometry to treat the pressure balance between the fluxtube and its surroundings, accounted for energy transport by radiation in the diffusion approximation, and treated convective energy transport with the mixing-length theory, taking into account the inhibition by the magnetic field in a simplified way.

The magnetohydrostatic equations have been numerically solved for the case of rotationally symmetric fluxtubes embedded in the field-free, stratified solar atmosphere (with a thin current sheet at the boundary between the fluxtube and the surrounding medium), without any restricting assumptions like the "thin-tube approximation" (Steiner et al., 1986). As however no energy equation was specified in these calculations, the internal temperature structure is simply specified as a free parameter (influencing the field geometry, opacity, etc.). Steiner and Pizzo (1989) have explored how the computed properties of the fluxtubes vary with the free model parameters, like the temperature or the fluxtube radius. The models can be used to compute synthetic Stokes spectra (e.g. Stokes V line ratios), which may be compared with the observed spectra to constrain the free model parameters by the observations.

These rotationally symmetric 2-D models are presently being extended by coupling energy transport self-consistently to the magnetohydrostatic problem. An exploratory investigation of radiative transport for optically thin fluxtubes using grey opacities has been carried out by Kalkofen et al. (1986). Recently the more general problem of coupling 2-D radiative transfer to fully self-consistent magnetohydrostatic models, using proper solar opacities, has been solved (Steiner and Stenflo, 1989), and an exploratory treatment of convective energy transport has been made (Hess, 1988). When an energy equation is included self-consistently, the internal temperature structure is no more a free parameter, but is a consequence of the energy exchange and heating processes.

The main long-term objective of such fluxtube modelling is to empirically constrain the heating processes due to dissipation of wave energy (cf. Ulmschneider and Muchmore, 1986), and to determine the spatial distribution of the wave heating. The general approach to accomplish this is the following: Self-consistent

fluxtube models with radiative and convective energy transport, but without any wave heating terms in the energy equation, are constructed and used to compute synthetic Stokes line profiles. From the discrepancy between the observed and synthetic Stokes profiles constraints on the missing terms in the energy equation may be found. As a further step, such terms may be introduced in a parameterized form, so that the parameter values can be determined by an inversion technique involving iterative least squares fitting between observed and synthetic spectra. Similar inversions have been used to determine the empirical fluxtube temperature structure (Keller, 1988; cf. Sect. 3.4).

The most comprehensive 2-D fluxtube models so far are those developed in slab geometry (no variation of the physical parameters in one horizontal direction) by Deinzer et al. (1984 a, b), Knölker et al. (1987), Knölker and Schüssler (1988), and Grossmann-Doerth et al. (1988). The time-dependent MHD equations are solved in a compressible medium, with the aim of exploring the properties of the concentrated stationary end state of the "fluxtube" (in this case actually "flux slab"). Convection and radiative transport have been treated in a similar way as by Spruit (1976), except for the latest paper in the series (Grossmann-Doerth et al., 1988), in which the diffusion approximation has been replaced by a full radiative-transfer treatment, although for a grey atmosphere.

Fig. 9, from Grossmann-Doerth et al. (1988), serves to illustrate the state of the art of these models at the same time as it summarizes the main characteristics of fluxtubes, for the case when the slab thickness is about 150 km at continuum optical depth $\tau_{5000} = 1$ inside the flux slab. The model presented has approached, but not yet quite reached its final steady state. Due to the partial evacuation of the fluxtube interior, the density contours are depressed inside the slab (Wilson depression). Accordingly, the level of $\tau_{5000} = 1$ in the external atmosphere (serving as the zero point of the geometrical height scale) is located about 180 km higher than the corresponding continuum level inside the flux slab. In the upper photosphere the field is expanding rapidly with height due to the decreasing external gas pressure, and magnetic tension forces are important.

The depression of the temperature contours in the lower, subsurface portion of the diagram is due to inhibition of lateral convective energy transport by the magnetic field, leading to cooling of the flux slab. The elevation of the temperature contours in the upper, observable layers is an effect of the 2-D radiative energy transfer. The gas in these layers is reached and heated by radiation coming from the hot bottom and walls of the flux slab, since the slab interior becomes fairly transparent due to the partial evacuation required by pressure balance.

This 2-D radiative transfer effect may be the main reason why small fluxtubes, in the quiet photospheric network or in active-region faculae or plages, appear as bright features in photospheric radiation (in the upper photosphere and chromosphere, wave heating may also be a contributing effect). For fluxtubes with larger diameters, the horizontal optical thickness insulates the interior more from radiative exchange, and the cooling due to inhibition of convection may dominate. They therefore appear as darker structures (pores or sunspots). The transition between bright and dark structures occurs for a slab thickness of about 500 km (Knölker and Schüssler, 1988).

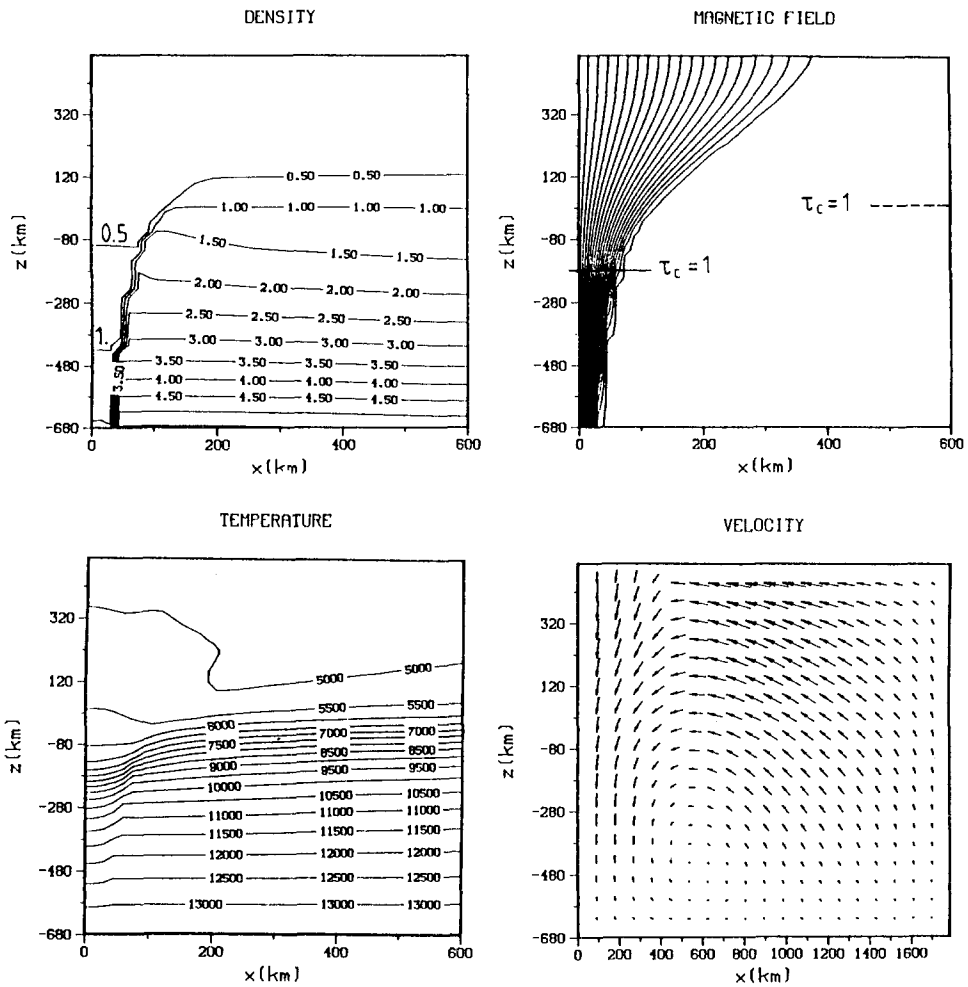

Fig. 9. Slab model for solar magnetic elements (Grossmann-Doerth et al., 1988). Since the model is symmetric around the center of the slab, only half of it is shown, and the horizontal coordinate scale has been stretched to enhance the visibility of the structure. The densities are in units of 3×10^{-7} g cm^{-3}, the temperatures in K

5.3. Origin inside the sun

Cosmic magnetic fields, observed in galaxies, stars, or planets, are believed to be generated by "dynamo" processes through the amplification of arbitrarily weak but finite "seed" magnetic fields as they interact with a turbulent, rotating, and electrically conducting medium (Parker, 1955; Zeldovich et al., 1983). The Coriolis forces induced by the rotation make the turbulence "cyclonic" and breaks the symmetry between left and right. Since this symmetry breaking is in the same sense over each hemisphere of the rotating medium, amplification of large-scale, mean magnetic fields is possible by the integrated, net effect of the small-scale, turbulent motions. As the evolution of the sun's magnetic field can be observed

in great detail, the solar dynamo may serve as a testing ground for the dynamo theories.

The sun represents an oscillating dynamo, with a period of 22 yr. Interaction between rotation and convection leads to radial and latitudinal gradients of the angular velocity (differential rotation). The rotational shear may generate a toroidal field from an initially poloidal one. Active regions and sunspots are formed when the strong, toroidal field breaks through the solar surface. Cyclonic turbulence acting on the toroidal field induces electric currents along the field, corresponding to regenerated poloidal field components. Turbulent transport spreads this new poloidal field, which has a polarity opposite to the previous poloidal field, making it a global field that cancels and reverses the direction of the previous one. The time scale for this sequence is 11 yr, half the full magnetic cycle. In this scenario, originally due to Parker (1955), the solar dynamo can be regarded as an oscillation between a poloidal and toroidal field.

Two approaches have been used for a detailed modelling of the solar dynamo, the kinematic and the dynamic approach. In the kinematic dynamos the magnetic field plays a passive role without back reaction on the fluid motions, and the basic processes of differential rotation, turbulent diffusion, and dynamo amplification (often described by the so-called α-effect) are characterized by free parameters chosen independently of each other. Generally these kinematic dynamos are based on the use of a "mean-field electrodynamics" (Steenbeck and Krause, 1969; Krause and Rädler, 1980) to describe the average effects of the turbulent fluctuations. In the dynamic approach on the other hand, there are essentially no free parameters, but the MHD equations for the solar convection zone with rotation are solved numerically with a powerful computer (cf. Gilman, 1983). The efficiencies of field amplification, differential rotation, and turbulent diffusion cannot be chosen independently of each other, but are calculated selfconsistently as joint consequences of the interaction between convection and rotation.

The kinematic dynamos have been apparently successful by being able to reproduce in particular the observed latitude migration of the zones of solar activity (e.g. Steenbeck and Krause, 1969), but this is hardly surprising since the model contains so many free, adjustable parameters. The physically more correct dynamical models, on the other hand, having essentially no free parameters, lead to dynamos looking very different from that of the real sun (the dynamo of the numerical simulations for instance runs much too fast). One major reason for this discrepancy is the inadequate spatial resolution in the grid used to model the MHD effects of convection (due to limited computer resources). We expect the magnetic field in the convection zone to be highly intermittent on a small scale (similar to the situation at the solar surface), due to convective expulsion of the flux to the boundaries of the turbulent eddies, and due to flux fragmentation caused by a spectrum of turbulence that extends down to very small scales, beyond what can be resolved in foreseeable "brute-force" numerical models. If the magnetic "filling factor" is small, much of the convective motions will not directly contribute to dynamo amplification, leading to a reduced overall dynamo efficiency (cf. Schüssler, 1983, for a discussion of fluxtube dynamos).

Another fundamental problem has been to prevent the generally buoyant flux ropes from floating to the surface before the dynamo processes have had sufficient time to act on them. The only place where long-term storage of magnetic flux is possible appears to be at the bottom of the convection zone or in the convectively stable layers below it (Spiegel and Weiss, 1980), although this layer may be subject to disruption by the Rayleigh-Taylor instability (Cattaneo and Hughes, 1987). This has led to the development of boundary-layer dynamos, operating at the bottom of the convection zone (DeLuca, 1987), where the magnetic flux can be stored. Spontaneous deformations of the flux layer (by convective buffeting or by the Rayleigh-Taylor instability) causes flux to "break loose" and float to the surface, appearing as new bipolar magnetic flux replenishing the old surface magnetic field pattern. New mechanisms for the emission and submergence of flux have recently been proposed by Parker (1988 a, b, c).

The emergence of new flux can easily be observed if it comes in large quantities, leading to the formation of new, low-latitude active regions with sunspots. About two orders of magnitude more flux is however brought up in the form of smaller bipolar magnetic regions, so-called "ephemeral active regions" (Harvey and Martin, 1973; Harvey et al., 1975; Martin and Harvey, 1979), which have a much broader latitude distribution than ordinary active regions. Another two orders of magnitude more flux (as compared with the ephemeral active regions) emerges in the form of the ubiquitous so-called intranetwork elements (Zirin, 1987), bipolar regions with sizes close to the resolution limit. As most of the flux emergence thus occurs on very small scales, much of it possibly beyond the resolution limit, the individual events will generally not be seen, but only the collective result of many emerged small-scale fluxtubes to form a new, regenerated flux *pattern*.

Indirect and unexpected evidence that this pattern is regenerated over a very short time scale, of the order of weeks, comes from observations of the differential rotation rate of the magnetic-field pattern. Two different methods have been used to determine the rotational phase velocity of the magnetic-field pattern recorded in full-disk magnetograms, with vastly different results. In one of the two methods, the longitude displacement (proper motion) of the field pattern between successive full-disk magnetograms is determined by cross-correlation analysis (Snodgrass, 1983). The magnetograms cannot be separated in time by more than 4 days due to foreshortening effects when a given pattern approaches the solar limb. The most reliable results are obtained for a lag of about 2 days (Snodgrass, 1983). The other method is based on an autocorrelation analysis of time series obtained by sampling the magnetic field at the central meridian in the successive magnetograms (Wilcox and Howard, 1970; Wilcox et al., 1970; Stenflo, 1974, 1989). Peaks in the autocorrelation functions are obtained at lags corresponding to an integer number of solar rotation periods, as the field pattern recurs at the central meridian after each rotation (provided that the pattern lifetime exceeds 27 days, which is the case).

Figure 10, from Stenflo (1989), summarizes the results. The solid curve has been obtained from the cross-correlation analysis of Snodgrass (1983), based on the 2-day lag. It agrees quite well with the dashed-dotted curve, which has

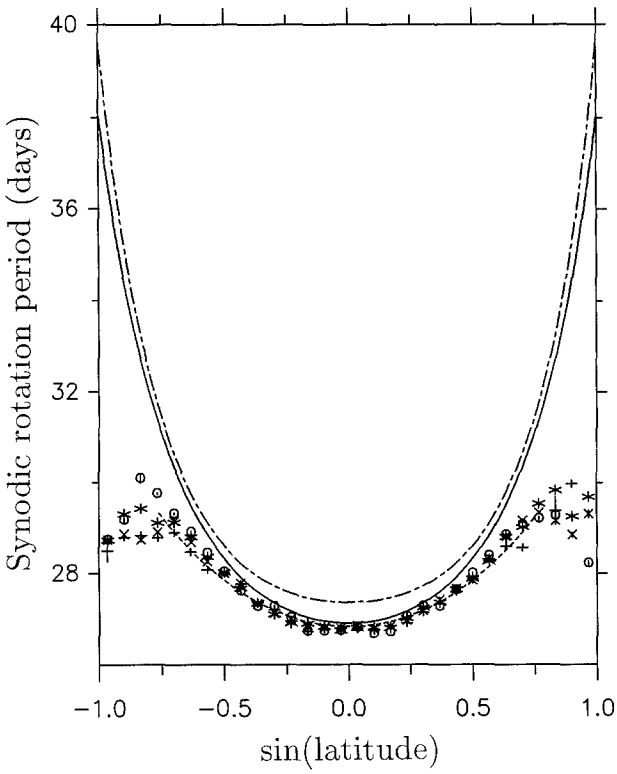

Fig. 10. Differential rotation of the sun's magnetic-field pattern, from Stenflo (1989). The solid curve gives the magnetic-field rotation law of Snodgrass (1983), based on cross-correlation analysis with a lag of two days. It agrees well with the Doppler rotation law (dashed-dotted), from Howard et al. (1983). The four different symbols represent the magnetic-field rotation law that is obtained from an autocorrelation analysis with lags of one (open circles), two (stars), three (crosses), and four (pluses) rotation periods. The dashed curve "hidden" among the symbols was obtained in an earlier autocorrelation analysis using a smaller data base (Stenflo, 1974). It is argued that the rotation law of the autocorrelation analysis represents the rotation rate of the *sources* of magnetic flux inside the sun, while that of Snodgrass represents the *surface* rotation rate

been obtained through Doppler measurements of the rotational plasma velocity (Howard et al., 1983). Over time scales not exceeding a few days, the phase velocity of the magnetic pattern is thus similar to the plasma velocity in the photosphere (the derived phase velocity is a statistical average of 15 yr of data). This is to be contrasted with the autocorrelation results, based on time scales or lags being a multiple of the rotation period (27–29 days). The open circles represent the results from the first autocorrelation peak, the stars the second, the crosses the third, and the pluses the fourth peak (lag of four times the rotation period), based on analysis of 26 yr long time series (Stenflo, 1989). The dashed curve among the symbols represents previous results using the same method but a much smaller data base (Stenflo, 1974).

While the rotation period of the Snodgrass (1983) curve increases steeply to 38 days at the poles, the autocorrelation periods reach a maximum of 29–30 days at

a latitude of 50–55°, and then decrease slightly towards the poles (polar spin-up). This huge discrepancy between the results of the two methods has nothing to do with instrumental or analysis errors, but reveals an intrinsic property of the sun, which we need to identify next.

If we for a moment ignore the Snodgrass curve, the quasi-rigid rotation of the autocorrelation results with respect to the Doppler curve can easily be explained in terms of surface redistribution (by 2-D turbulent diffusion, rotational shearing, and meridional circulation) of old, remnant magnetic fluxes, as demonstrated in great detail through numerical simulations by Sheeley et al. (1987). The flux pattern, which initially has a phase velocity coinciding with the plasma velocity, develops through these redistribution processes a large-scale pattern with a phase velocity that becomes increasingly rigid. A cross-correlation analysis picks up the pattern phase velocity, which is quasi-rigid and greatly different from the strongly differential plasma rotation rate (Doppler curve).

The reason why this scenario is inconsistent with the behaviour of the real sun is that it cannot explain why the Snodgrass phase velocity agrees so well with the plasma velocity. In the Sheeley et al. (1987) model it does not make much difference if lags of 2 days or 27 days are used to determine the pattern phase velocity through cross correlation. In both cases quasi-rigid rotation is predicted by the model. Surface flux redistribution is thus unable to explain the large lag dependence of the rotation curves.

The only possibility of explaining the coexistence of the two pattern phase velocities appears to be if the pattern is constantly being replenished over a time scale of weeks (larger than a few days, but shorter than the rotation period) from sources in the solar interior, and that the rotational phase velocity of the disturbances in the source region is different and more rigid as compared with the plasma velocity in the photosphere. In this scenario there must be agreement between the pattern phase velocity and the plasma velocity on short time scales, as Snodgrass (1983) finds, since the pattern that we see at the surface at any given time is made up of newly emerged fields, not of redistributed old flux (cf. Martin, 1988). Since however this pattern is being replaced by a new one within one rotation period, the peaks in the autocorrelation analysis do not represent correlations between the *same* fluxes, but between "active longitudes" of flux emergence. These "active longitudes" rotate more rigidly, being representative of the rotation rate of the deep-seated sources of flux. Since the most likely location of these sources of flux is at the bottom of the convection zone, the observed surface rotation laws provide constraints on the depth variation of the angular velocity of rotation (Stenflo, 1989). Since the individual flux emergence events probably occur on very small space and time scales, only their collective pattern effects are resolved.

Other clues to the origin of solar magnetic fields come from a harmonic mode analysis of the evolution of the surface magnetic field pattern. If the evolution of the magnetic fields throughout the sun can be described in terms of some linear global wave equation having the spherical harmonics as its eigenfunctions (as is generally the case in the kinematic dynamo approach, cf. Steenbeck and Krause, 1969), then a time series analysis of the harmonic coefficients would reveal the

eigenmode frequencies and thereby constrain the properties of the underlying wave equation.

Using a 26 yr data base of Mt Wilson and Kitt Peak magnetograms, decomposing the observed field pattern in its spherical harmonic components, and performing a power spectrum analysis on the time series of the harmonic coefficients, a set of discrete, resonant frequencies was indeed uncovered (Stenflo and Vogel, 1986; Stenflo and Güdel, 1988; Stenflo, 1988). The harmonic modes can be characterized by their degree ℓ and order m. The rotationally symmetric modes (with $m = 0$) are found to strictly obey a parity selection rule. While modes with odd parity (odd values of ℓ, corresponding to patterns anti-symmetric with respect to reflections in the equatorial plane) are dominated by power sharply centered at the frequency $(22\ \mathrm{yr})^{-1}$, with less than 10 % of the power at the second harmonic at $(11\ \mathrm{yr})^{-1}$, the even parity modes (corresponding to patterns symmetric with respect to reflections in the equatorial plane) show no trace of the 22 yr cycle. Instead they reveal weaker but still discrete power spectrum peaks for which the frequency increases with the value of ℓ, reminiscent of the way in which the frequencies of the global acoustic modes (in the 5 min range) increase with ℓ. This sequence of peaks has maximum power at a period of about 2 yr for $\ell = 6$. Various interpretations of these modes have been provided, in terms of stochastic excitation of dynamo modes (Hoyng, 1987), or in terms of global magnetoacoustic waves, trapped in a cavity between the toroidal field belts in the northern and southern hemispheres (Brandenburg, 1986). The non-axisymmetric modes (with $m \neq 0$) are dominated by the 22 yr resonance but exhibit identical behaviour for odd and even parity (odd and even $\ell - m$). In addition, a number of higher discrete frequencies with smaller amplitudes are revealed (Stenflo and Güdel, 1988).

Helioseismology will be an important future tool for uncovering the secrets of the solar dynamo. Inversion of the observed rotational splittings of the p-mode oscillations will lead to an empirical determination of the depth and latitude variations of the sun's rotation rate (Gough, 1985; Duvall et al., 1987; Brown and Morrow, 1987), key parameters in the dynamo problem. There is also increasing hope that helioseismology may allow the spatial distribution of magnetic flux inside the sun to be determined. We have given arguments above that the magnetic flux inside the sun is likely to exist in highly intermittent form. The magnetic-field inhomogeneities scatter sound waves and collectively cause shifts of the resonant frequencies of the global p modes. The theory for calculating such shifts is being developed in a recent series of papers (Bogdan and Zweibel, 1987; Bogdan, 1987; Bogdan and Cattaneo, 1989), with the ultimate aim of diagnosing the magnetic field distribution inside the sun.

6. Concluding remarks

We have seen that the magnetic flux on the solar surface (photosphere) occurs in the form of discrete structures (which we call "fluxtubes"), which except for the sunspots are smaller than what can be spatially resolved with present-day telescopes. The resolved flux patterns that we can study directly in magnetograms refer either to clusters of fluxtubes, or to sunspots (although sunspots also contain complex internal fine structures that are far from being fully resolved). These circumstances have made it necessary to rely heavily on indirect diagnostic methods, using polarized line profiles or multi-line information, to overcome the resolution limits. Although a wealth of data on the internal fluxtube properties have been gained this way, this success has come at the expense of giving up information on evolution and morphology.

The key to a unified understanding of the physics of the solar atmosphere, like solar activity, atmospheric heating, etc., may be found in the MHD processes of the basic magnetic elements. The processes relating to the origin and evolution of these elements, like convective collapse of magnetic flux, fluting instability, reconnection and flux disappearance, waves, etc., generally take place on space and time scales beyond present resolution limits.

The future prospects of directly observing these crucially important physical processes are however quite promising. There are good reasons why the size spectrum of the magnetic elements does not continue indefinitely towards smaller and smaller sizes, but that there is an approximate lower limit, of the order of 100 km, around which much of the strong-field flux accumulates (cf. Sect. 2.3, 5.1, and 3.6 above). As future solar telescopes in an already advanced planning stage, to be used from space (OSL, "Orbiting Solar Laboratory") or from ground (LEST, "Large Earth-based Solar Telescope") aim at observing magnetic fields with a resolution that can reach 0.1 arcsec ($\approx$ 73 km), the basic fluxtubes are within reach. The observational constraints (presented in Sect. 4) on the properties of the "turbulent" magnetic fluxes and weaker fields between the strong-field fluxtubes indicate that we are now close to a positive detection of these "hidden" fields. Future access to this still unexplored territory opens the possibility of direct investigation of the basic interactions between the magnetic fields and the turbulent solar plasma.

It is thus no wonder that the primary objective of major new telescopes is to observe the small-scale solar magnetic field, and that high spatial resolution is combined with high spectral resolution and simultaneous wavelength coverage, in the recognition that direct resolution must be accompanied by indirect spectral diagnostics to allow a quantitative analysis of the small-scale magnetic structures. In the near future intermediate-size, national projects, like the US ASP (Advanced Stokes Polarimeter) and the French THEMIS, will provide new tools of this kind. This will be followed by LEST, a next-generation type facility with a 2.4 m, "polarization-free" telescope that will be helium filled to reduce the internal "seeing" effects, and will make use of a number of new technologies, like adaptive optics, to overcome limitations by the earth's atmosphere (Stenflo and Engvold, 1988). As an international framework is needed for a project of this magnitude,

an international organization, the LEST Foundation, was established in 1983 to carry out this ambitious venture. 10 nations are presently participating in the realization of LEST.

These joint efforts on a global scale, together with the recent advances in observational methods, interpretation, and theoretical understanding of the small-scale solar magnetic fields, indicate the rich future potential for exploring the plasma physics of the sun. Both the methods developed and the knowledge acquired are applicable to other stars as well. We are beginning to unravel the fundamental plasma physics processes in the solar atmosphere, and seem to be on the verge of a unified understanding of solar and stellar activity.

Acknowledgements. I am grateful to Tom Bogdan for his detailed comments on the draft of this review, and to the High Altitude Observatory for its hospitality during my sabbatical year.

References

Alfvén, H., Fälthammar, C.-G.: 1963, *Cosmical Electrodynamics — Fundamental Principles, 2nd Ed.*, Oxford Univ. Press
Altschuler, M.D., Newkirk, G.N.Jr.: 1969, *Solar Phys.* **9**, 131
Ayres, T.R., Testerman, L., Brault, J.W.: 1986, *Astrophys. J.* **304**, 542
Bogdan, T.J.: 1987, *Astrophys. J.* **318**, 888
Bogdan, T.J., Cattaneo, F.: 1989, *Astrophys. J.* **342**, in press
Bogdan, T.J., Zweibel, E.G.: 1987, *Astrophys. J.* **312**, 444
Bogdan, T.J., Gilman, P.A., Lerche, I., Howard, R.: 1988, *Astrophys. J.* **327**, 451
Brandenburg, A.: 1986, *Proc. 7th Finnish-Soviet Coll.*, Tallinn, December 1986
Brandt, P.N., Scharmer, G.B., Ferguson, S., Shine, R.A., Tarbell, T.D., Title, A.M.: 1988, *Nature* **335**, 238
Brault, J.W., Noyes, R.: 1983, *Astrophys. J.* **269**, L61
Brown, T.M., Morrow, C.A.: 1987, *Astrophys. J.* **314**, L21
Cattaneo, F., Hughes, D.W.: 1987, in G. Athay, D.S. Spicer (eds.), *Theoretical Problems in High Resolution Solar Physics II*, NASA Conf. Publ. 2483, p. 101
Chandrasekhar, S.: 1950, *Radiative Transfer*, Clarendon Press, Oxford
Chapman, G.A., Sheeley, N.R.Jr.: 1968, *Solar Phys.* **5**, 442
Deinzer, W., Hensler, G., Schüssler, M., Weisshaar, E.: 1984 a, *Astron. Astrophys.* **139**, 426
Deinzer, W., Hensler, G., Schüssler, M., Weisshaar, E.: 1984 b, *Astron. Astrophys.* **139**, 435
DeLuca, E.E.: 1987, Ph. D. Thesis, Univ. Colorado (available as NCAR Cooperative Thesis No. 104)
Deming, D., Boyle, R.J., Jennings, D.E., Wiedemann, G.: 1988, *Astrophys. J.* **333**, 978
Deubner, F.-L., Göhring, R.: 1970, *Solar Phys.* **13**, 118
Duvall, T.L. Jr., Harvey, J.W., Pomerantz, M.A.: 1987, *Astrophys. Space Sci.* **137**, 19
Frazier, E.N.: 1970, *Solar Phys.* **14**, 89
Frazier, E.N., Stenflo, J.O.: 1972, *Solar Phys.* **27**, 330
Frazier, E.N., Stenflo, J.O.: 1978, *Astron. Astrophys.* **70**, 789
Gilman, P.A.: 1983, in J.O. Stenflo (ed.), *Solar and Stellar Magnetic Fields: Origins and Coronal Effects, IAU Symp.* **102**, 247
Giovanelli, R.G.: 1980, *Solar Phys.* **68**, 49
Giovanelli, R.G.: 1982, *Solar Phys.* **80**, 21
Gokhale, M.H., Zwaan, C.: 1972, *Solar Phys.* **26**, 52
Gough, D.: 1985, *Solar Phys.* **100**, 65
Grossmann-Doerth, U., Pahlke, K.-D., Schüssler, M.: 1987, *Astron. Astrophys.* **176**, 139
Grossmann-Doerth, U., Knölker, M., Schüssler, M., Weisshaar, E.: 1988, in *Solar and Stellar Granulation*, Proc. Third International Workshop of the OAC — NATO Advanced Research Workshop, Capri, Italy, June 21-25, 1988

Grossmann-Doerth, U., Schüssler, M., Solanki, S.K.: 1988, *Astron. Astrophys.* **206**, L37
Hale, G.E.: 1908, *Astrophys. J.* **28**, 100
Harvey, J.W.: 1977, in E.A. Müller (ed.), *Highlights of Astronomy* **4**, 223
Harvey, J.W.: 1986, in W. Deinzer, M. Knölker, H.H. Voigt (eds.), *Small Scale Magnetic Flux Concentrations in the Solar Photosphere*, Vandenhoeck & Ruprecht, Göttingen, p. 25
Harvey, K.L., Harvey, J.W.: 1973, *Solar Phys.* **28**, 61
Harvey, K.L., Martin, S.F.: 1973, *Solar Phys.* **32**, 389
Harvey, K.L., Harvey, J.W., Martin, S.F.: 1975, *Solar Phys.* **40**, 87
Hasan, S.S.: 1984, *Astrophys. J* **285**, 851
Hasan, S.S.: 1985, *Astron. Astrophys.* **143**, 39
Hess, R.: 1988, *Diplomarbeit*, Inst. Astron., ETH Zürich
Howard, R., Stenflo, J.O.: 1972, *Solar Phys.* **22**, 402
Howard, R., Adkins, J.M., Boyden, J.E., Cragg, T.A., Gregory, T.S., LaBonte, B.J., Padilla, S.P., Webster, L.: 1983, *Solar Phys.* **83**, 321
Hoyng, P.: 1987, *Astron. Astrophys.* **171**, 357
Illing, R.M.E., Landman, D.A., Mickey, D.L.: 1975, *Astron. Astrophys.* **41**, 183
Jones, H.P., Giovanelli, R.G.: 1983, *Solar Phys.* **87**, 37
Kalkofen, W., Rosner, R., Ferrari, A., Massaglia, S.: 1986, *Astrophys. J.* **304**, 519
Keller, C.: 1988, *Diplomarbeit*, Inst. Astron., ETH Zürich
Kneer, F.: 1973, *Solar Phys.* **28**, 361
Knölker, M., Schüssler, M.: 1988, *Astron. Astrophys.* **202**, 275
Knölker, M., Schüssler, M., Weisshaar, E.: 1988, *Astron. Astrophys.* **194**, 257
Krause, F., Rädler, K.-H.: 1980, *Mean-Field Magnetohydrodynamics*, Pergamon Press, Oxford
Lites, B.W., Scharmer, G.B.: 1988, in O. v.d. Lühe (ed.), *High Spatial Resolution Solar Observations*, Proc. Tenth Sacramento Peak Summer Workshop, Sunspot, NM, August 22-26, 1988
Lites,B.W., Skumanich, A.: 1989, *Astrophys. J.*, in press
Lites, B.W., Skumanich, A., Rees, D.E., Murphy, G.A., Carlsson, M.: 1987, *Astrophys. J.* **318**, 930
Lites, B.W., Skumanich, A., Rees, D.E., Murphy, G.A.: 1988, *Astrophys. J.* **330**, 493
Livi, S.H.B., Wang, J., Martin, S.F.: 1985, *Australian J. Phys.* **38**, 855
Martin, S.F.: 1988, *Solar Phys.* **117**, 243
Martin, S.F., Harvey, K.L.: 1979, *Solar Phys.* **64**, 93
Martin, S.F., Livi, S.H.B., Wang, J.: 1985 a, *Australian J. Phys.* **38**, 929
Martin, S.F., Livi, S.H.B., Wang, J., Shi, Z.: 1985 b, in M.J. Hagyard (ed.), *Measurements of Solar Vector Magnetic Fields*, NASA Conf. Publ. 2374, p. 403
Mathys, G., Solanki, S.K.: 1989, *Astron.Astrophys.*, in press
Mathys, G., Stenflo, J.O.: 1986, *Astron.Astrophys.* **168**, 184
McMath, R.R., Mohler, O.C., Pierce, A.K., Goldberg, L.: 1956, *Astrophys. J.* **124**, 1
Mehltretter, J.P.: 1974, *Solar Phys.* **38**, 43
Meyer, E., Schmidt, H.U., Weiss, N.O.: 1977, *Monthly Notices Roy. Astron. Soc.* **179**, 741
Nordlund, Å.: 1983, in J.O. Stenflo (ed.), *Solar and Stellar Magnetic Fields: Origins and Coronal Effects, IAU Symp.* **102**, 79
Nordlund, Å.: 1986, in W. Deinzer, M. Knölker, H.H. Voigt (eds.), *Small Scale Magnetic Flux Concentrations in the Solar Photosphere*, Vandenhoeck & Ruprecht, Göttingen, p. 83
Pantellini, F.G.E., Solanki, S.K., Stenflo, J.O.: 1988, *Astron. Astrophys.* **189**, 263
Parker, E.N.: 1955, *Astrophys. J.* **122**, 293
Parker, E.N.: 1978, *Astrophys. J.* **221**, 368
Parker, E.N.: 1979 a, *Cosmical Magnetic Fields*, Clarendon Press, Oxford
Parker, E.N.: 1979 b, *Astrophys. J.* **230**, 905
Parker, E.N.: 1988 a, *Astrophys. J.* **325**, 880
Parker, E.N.: 1988 b, *Astrophys. J.* **326**, 395
Parker, E.N.: 1988 c, *Astrophys. J.* **326**, 407
Peckover, R.S., Weiss, N.O.: 1972, *Computer Phys. Communications* **4**, 339
Sheeley, N.R.Jr., Nash, A.G., Wang, Y.-M.: 1987, *Astrophys. J.* **319**, 481
Schüssler, M.: 1983, in J.O. Stenflo (ed.), *Solar and Stellar Magnetic Fields: Origins and Coronal Effects, IAU Symp.* **102**, 213
Schüssler, M.: 1984, *Astron. Astrophys.* **140**, 453
Schüssler, M.: 1986, in W. Deinzer, M. Knölker, H.H. Voigt (eds.), *Small Scale Magnetic Flux Concentrations in the Solar Photosphere*, Vandenhoeck & Ruprecht, Göttingen, p. 103
Schüssler, M.: 1987, in E.H. Schröter, M. Vázquez, A.A. Wyller (eds.), *The Role of Fine-scale Magnetic Fields on the Structure of the Solar Atmosphere*, Cambridge Univ. Press, p. 223
Sheeley, N.R.Jr.: 1967, *Solar Phys.* **1**, 171

Simon, G.W., Leighton, R.B.: 1964, *Astrophys. J* **140**, 1120
Simon, G.W., Title, A.M., Topka, K.P., Tarbell, T.D., Shine, R.A., Ferguson, S.H., Zirin, H., SOUP Team: 1988, *Astrophys. J.* **327**, 964
Skumanich, A., Lites, B.W.: 1987, *Astrophys. J.* **322**, 473
Snodgrass, H.B.: 1983, *Astrophys. J.* **270**, 288
Solanki, S.K.: 1986, *Astron. Astrophys.* **168**, 311
Solanki, S.K.: 1987 a, in E.H. Schröter, M. Vázquez, A.A. Wyller (eds.), *The Role of Fine-scale Magnetic Fields on the Structure of the Solar Atmosphere*, Cambridge Univ. Press, p. 67
Solanki, S.K.: 1987 b, in L. Hejna, M. Sobotka (eds.), *Proc. Tenth European Regional Astronomy Meeting of the IAU. Vol. 1: The Sun*, Publ. Astron. Inst. Czechosl. Acad. Sci., p. 95
Solanki, S.K., Pahlke, K.D.: 1988, *Astron. Astrophys.* **201**, 143
Solanki, S.K., Steenbock, W.: 1988, *Astron. Astrophys.* **189**, 243
Solanki, S.K., Stenflo, J.O.: 1984, *Astron. Astrophys.* **140**, 185
Solanki, S.K., Stenflo, J.O.: 1985, *Astron. Astrophys.* **148**, 123
Spiegel, E.A., Weiss, N.O.: 1980, *Nature* **287**, 616
Spruit, H.C.: 1976, *Solar Phys.* **50**, 269
Spruit, H.C.: 1979, *Solar Phys.* **61**, 363
Spruit, H.C.: 1981, in L.E. Cram, J.H. Thomas (eds.), *The Physics of Sunspots*, NSO/Sacramento Peak, p. 98
Spruit, H.C.: 1983, in J.O. Stenflo (ed.), *Solar and Stellar Magnetic Fields: Origins and Coronal Effects, IAU Symp.* **102**, 41
Spruit, H.C., Roberts, B.: 1983, *Nature* **304**, 401
Spruit, H.C., Zwaan, C.: 1981, *Solar Phys.* **70**, 207
Spruit, H.C., Zweibel, E.G.: 1979, *Solar Phys.* **62**, 15
Spruit, H.C., Title, A.M., van Ballegooijen, A.A.: 1987, *Solar Phys.* **110**, 115
Steenbeck, M., Krause, F.: 1969, *Astron. Nachr.* **291**, 49
Steiner, O., Pizzo, V.J.: 1989, *Astron. Astrophys.* **211**, 447
Steiner, O., Stenflo, J.O.: 1989, *IAV Symp.* **138**, in press
Steiner, O., Pneuman, G.W., Stenflo, J.O.: 1986, *Astron. Astrophys.* **170**, 126
Stenflo, J.O.: 1968, *Acta Univ. Lund.* **II**, No. 2 = *Medd. Lund Astron. Observ.* **II**, No. 153
Stenflo, J.O.: 1973, *Solar Phys.* **32**, 41
Stenflo, J.O.: 1974, *Solar Phys.* **36**, 495
Stenflo, J.O.: 1975, *Solar Phys.* **42**, 79
Stenflo, J.O.: 1976, in V. Bumba, J. Kleczek (eds.), *Basic Mechanisms of Solar Activity, IAU Symp.* **71**, 69
Stenflo, J.O.: 1977, in R.M. Bonnet, P. Delache (eds.), *The Energy Balance and Hydrodynamics of the Solar Chromosphere and Corona, IAU Coll.* **36**, 143
Stenflo, J.O.: 1978, *Rep. Prog. Phys.* **41**, 865
Stenflo, J.O.: 1982, *Solar Phys.* **80**, 209
Stenflo, J.O.: 1984, *Adv. Space Res.* **4**, 5
Stenflo, J.O.: 1985, *Solar Phys.* **100**, 189
Stenflo, J.O.: 1986, *Mitt. Astron. Gesellschaft Nr. 65*, p. 25
Stenflo, J.O.: 1987, *Solar Phys.* **114**, 1
Stenflo, J.O.: 1988, *Astrophys. Space Sci.* **144**, 321
Stenflo, J.O.: 1989, *Astron. Astrophys.* **210**, 403
Stenflo, J.O., Engvold, O.: 1988, in O. v.d. Lühe (ed.), *High Spatial Resolution Solar Observations*, Proc. Tenth Sacramento Peak Summer Workshop, Sunspot, NM, August 22-26, 1988
Stenflo, J.O., Güdel, M.: 1988, *Astron. Astrophys.* **191**, 137
Stenflo, J.O., Harvey, J.W.: 1985, *Solar Phys.* **95**, 99
Stenflo, J.O., Lindegren, L.: 1977, *Astron. Astrophys.* **59**, 367
Stenflo, J.O., Vogel, M.: 1986, *Nature* **319**, 285
Stenflo, J.O., Baur, T.G., Elmore, D.F.: 1980, *Astron. Astrophys.* **84**, 60
Stenflo, J.O., Twerenbold, D., Harvey, J.W.: 1983 a, *Astron. Astrophys. Suppl. Ser.* **52**, 161
Stenflo, J.O., Twerenbold, D., Harvey, J.W., Brault, J.W.: 1983 b, *Astron. Astrophys. Suppl. Ser.* **54**, 505
Stenflo, J.O., Harvey, J.W., Brault, J.W., Solanki, S.K: 1984, *Astron. Astrophys.* **131**, 33
Stenflo, J.O., Solanki, S.K., Harvey, J.W.: 1987 a, *Astron. Astrophys.* **171**, 305
Stenflo, J.O., Solanki, S.K., Harvey, J.W.: 1987 b, *Astron. Astrophys.* **173**, 167
Stenholm, L.G., Stenflo, J.O.: 1978, *Astron. Astrophys.* **67**, 33
Tanenbaum, A.S., Wilcox, J.M., Frazier, E.N., Howard, R.: 1969, *Solar Phys.* **9**, 328

Tarbell, T.D., Title, A.M., Schoolman, S.A.: 1979, *Astrophys. J.* **229**, 387
Tarbell, T., Ferguson, S., Frank, Z., Title, A., Topka, K.: 1988, in O. v.d. Lühe (ed.), *High Spatial Resolution Solar Observations*, Proc. Tenth Sacramento Peak Summer Workshop, Sunspot, NM, August 22-26, 1988
Title, A.M., Tarbell, T.D., Topka, K.P.: 1987, *Astrophys. J.* **317**, 892
Topka, K., Ferguson, S., Frank, Z., Tarbell, T., Title, A.: 1988, in O. v.d. Lühe (ed.), *High Spatial Resolution Solar Observations*, Proc. Tenth Sacramento Peak Summer Workshop, Sunspot, NM, August 22-26, 1988
Ulmschneider, P., Muchmore, D.: 1986, in W. Deinzer, M. Knölker, H.H. Voigt (eds.), *Small Scale Magnetic Flux Concentrations in the Solar Photosphere*, Vandenhoeck & Ruprecht, Göttingen, p. 191
Unno, W., Ando, H.: 1979, *Geophys. Astrophys. Fluid Dyn.* **12**, 107
van Ballegooijen, A.A.: 1985, in M.J. Hagyard (ed.), *Measurements of Solar Vector Magnetic Fields*, NASA Conf. Publ. 2374, p. 322
Venkatakrishnan, P.: 1985, *J. Astrophys. Astron.* **6**, 21
Venkatakrishnan, P.: 1986, *Solar Phys.* **104**, 347
Wallenhorst, S.G., Howard, R.: 1982, *Solar Phys.* **76**, 203
Wallenhorst, S.G., Topka, K.P.: 1982, *Solar Phys.* **81**, 33
Wang, H.: 1988, *Solar Phys.* **116**, 1
Weiss, N.O.: 1966, *Proc. Royal Soc. London A* **293**, 310
Wilcox, J.M., Howard, R.: 1970, *Solar Phys.* **13**, 251
Wilcox, J.M., Schatten, K.H., Tanenbaum, A.S., Howard, R.: 1970, *Solar Phys.* **14**, 255
Wittmann, A.: 1974, *Solar Phys.* **36**, 29
Zayer, I., Solanki, S.K., Stenflo, J.O.: 1989, *Astron. Astrophys.* **211**, 463
Zeldovich, Ya.B., Ruzmaikin, A.A., Sokoloff, D.D.: 1983, *Magnetic Fields in Astrophysics*, Gordon and Breach
Zirin, H.: 1987, *Solar Phys.* **110**, 101
Zirin, H., Popp, B.: 1989, *Astrophys. J.*, in press
Zwaan, C.: 1978, *Solar Phys.* **60**, 213
Zwaan, C.: 1985, *Solar Phys.* **100**, 397
Zwaan, C.: 1987, *Ann. Rev. Astron. Astrophys.* **25**, 83
Zwaan, C., Brants, J.J., Cram, L.E.: 1985, *Solar Phys.* **95**, 3

Received December 1, 1988; accepted January 30, 1989

This article was processed by the author
using the Springer-Verlag TEX Theaar macro package 1988.