REMOTE SENSING OF SOLAR MAGNETIC FIELDS

Bruce W. Lites
High Altitude Observatory
National Center for Atmospheric Research
Boulder, Colorado

Abstract. New techniques for remote sensing of solar magnetic fields now provide measures of the magnetic field vector within the solar atmosphere with high angular resolution and high precision. These measurements have enabled a much improved physical understanding of magnetic processes and phenomena in the solar atmosphere, processes that drive the variability of the Sun's radiative and particulate output. The new techniques are reviewed here in the context of the scientific advances they have fostered. Emphasis is given to techniques for inferring the field vector. The quantitative nature of the information needed to explore the solar phenomena sharply constrains the needed precision and angular resolution of the observations. These requirements are reviewed here, along with an assessment of how future improvements in observing capabilities will address these requirements. One may also attribute much of the recent advance in our understanding of solar magnetic fields to ongoing progress in techniques for analysis of the polarization measurements that underlie solar magnetometry. The status and prospects of analysis techniques are also reviewed.

1. SOLAR MAGNETISM AND SOLAR VARIABILITY

The magnetic field of the Sun varies on a wide range of timescales and spatial scales. As a consequence, we experience dramatic variations in the Earth's space environment. Even the total energy output of the Sun varies measurably (~0.1% [Fröhlich and Lean, 1998]) under the influence of the variable solar magnetic field, probably as a result of its influence on the visible and near-infrared radiation emitted by the surface layers of the solar atmosphere (the photosphere). The Sun is slightly brighter near the maximum of the 11-year sunspot activity cycle. Even larger fluctuations may occur on timescales of centuries or longer. In its "quiet" state the 6000-K photosphere is enveloped by an extended, tenuous corona (Figure 1) having temperatures in excess of $10^6$ K, causing the Sun to emit strongly in X rays. Magnetic fields play a crucial role in this remarkable but as yet not fully understood heating [Parker, 1987]. As the magnetic fields of solar active regions erupt into the solar atmosphere, an array of fascinating and energetic phenomena transpire to produce variability of the solar output: Large fluctuations occur in the ultraviolet and X-ray output, solar flares (Figure 2) produce very rapid changes at short wavelengths that are sometimes accompanied by bursts of energetic protons at the Earth, and intrinsically magnetic events known as coronal mass ejections, or CMEs (Plate 1), may disrupt the interplanetary magnetic field, causing geomagnetic storms. Societal consequences of solar magnetic activity are widespread, including disruption of communications, radiation hazards for astronauts and for unmanned spacecraft [Baker et al., 1998], and likely influences on terrestrial climate [Frieman et al., 1994].

1.1. The Large-Scale Solar Magnetic Field and the Solar Dynamo

The 11-year cycle of solar activity is certainly the most widely recognized mode of solar variability. In fact, it represents a 22-year magnetic cycle: Two 11-year cycles are required to restore the dominant magnetic polarity near the poles of the Sun. This cycle is dramatically revealed in the synoptic "magnetic butterfly diagram" of Figure 3, which shows the evolution of the large-scale solar magnetic field during the past two sunspot cycles. This map is assembled from individual solar "magnetograms" of the full solar disk similar to that shown in Figure 4. The behavior of the solar magnetic field at these global scales reflects the working of the solar dynamo [Roberts, 1994]. The dynamo, which is believed to be operative near the base of the solar convective layer (or about one third the solar radius below the surface), results from the combined action of solar rotation and convective flows within the highly conducting solar plasma. The solar dynamo problem represents one of the major challenges of astrophysics: Dynamos are responsible for magnetic activity in a range of stellar types, and some stars sustain magnetic activity cycles that are much more energetic than that of the Sun. Figure 3 contains a wealth of information on the nature of the dynamo as reflected in the behavior of the magnetic fields as they penetrate the solar surface, for example, the latitude and phase in the cycle of the emergence of fields and the motions of large-scale unipolar
patches of field. Figure 3 thus represents an observational benchmark which any successful dynamo theory must reproduce in detail. We shall see that the observational data represented by this synoptic map contain only limited quantitative information regarding the magnetic fields themselves. Nonetheless, the value of long-term synoptic data of this sort is clearly evident.

1.2. Solar Active Regions and Sunspots

Individual active regions produce most of the variability of the Sun on timescales of weeks to months. Magnetic fields from one such active region are shown in enlarged view in Figure 4. The magnetic fields that comprise active regions are believed to be individual bundles of magnetic flux separated from the main dynamo region, which rise through the entire solar convective envelope in a few weeks or months. Usually, they present a dipolar field structure oriented roughly in the east-west direction, with leading polarities opposite in the Northern and Southern Hemispheres, as seen in Figure 4.

Magnetic buoyancy plays a dominant role in the evolution of solar magnetic fields. The magnetic field itself exerts pressure, so that the total pressure at any point in the atmosphere is the sum of the kinetic pressure of the gas plus the magnetic pressure. A flux rope (so-called because the magnetic lines of force follow twisted paths) within the solar atmosphere has much higher magnetic field strength than its nearly field-free surroundings, but
if it is in rough pressure equilibrium with its surroundings, the gas pressure within the rope must be reduced relative to that of its surroundings, and hence the rope will have a reduced density. The loop is therefore buoyant, and it will attempt to rise. Only viscosity of the surrounding unmagnetized plasma and the magnetic tension of the rope itself resist its inevitable rise toward the surface and beyond. (Magnetic fields also exert a tension force that tends to straighten the lines of force.) Indeed, the movies of the solar corona in X rays recorded by the Yohkoh spacecraft [Uchida et al., 1992] reveal a nearly ubiquitous expansion of the magnetized structures of the corona. Interaction of the rising flux with preexisting fields via magnetic reconnection [Forbes, 1992; Yokoyama and Shibata, 1995] can cause energetic solar flares such as that of the X-ray image of Figure 2. Some fraction of this magnetic flux escapes into interplanetary space via the CME process, some

Figure 2. This image, taken in a wavelength band around 171 Å by the Transition Region and Coronal Explorer (TRACE) on September 30, 1998, at 1421:23 UT, reveals postflare loops. The brightening near the tops of the loops may indicate the site of magnetic reconnection. There are both emitting and absorbing loops, and the erupting prominence in the foreground obscures the loops beyond. The repeating image pattern from the bright tops of the flare loops is due to diffraction from the fine mesh supporting the entrance filter of the telescope.
Plate 1. A coronal mass ejection (CME) observed by the large-angle spectroscopic C2 coronagraph (LASCO) aboard SOHO shows the large-scale disruption of the corona and the subsequent eruption of the underlying prominence. The frames at 10:05 and 10:29 UT reveal a shell of coronal material erupting outward, followed by the filamentary cool prominence material in subsequent frames. The rarefied region just above the prominence is the prominence cavity, thought to be the location of the strong, twisted magnetic fields which supported the prominence before eruption. CMEs occur about once per day, on average, releasing an energy of about $10^{31}$ ergs into interplanetary space, i.e., some $10^4$ times the Jovian impact energy of Shoemaker-Levy 9 comet fragments.
fraction is dissipated via turbulence of the solar atmosphere, and surviving flux diffuses and is forced to migrate poleward as a result of convective and large-scale motions in the solar envelope. The latter process may be seen clearly in Figure 3 where light or dark structures of the remnants of active regions move rapidly poleward during the declining phase of a solar cycle.

Convective flows also play an important role in the evolution of active regions. Convection on the scale of solar granulation confines most of the magnetic flux of active regions to the dark "intergranular lanes": zones of subsiding, cooler plasma between the hot, bright, upwelling granules themselves. The convective flows are believed to initiate the process of convective collapse of weak magnetic flux, leading to the kilogauss fields of tiny flux tubes [Parker, 1978]. At larger scales the convective flows at the supergranular scale, and also at scales intermediate between granules and supergranules (the so-called "mesogranulation"), structure plages and clearly participate in the dispersal of active-region flux. (Plages are strongly magnetized areas within active regions but outside of sunspots. They are composed mainly of small kilogauss flux tubes. They appear slightly brighter than the surrounding photosphere when viewed near the solar limb.) Where plage fields are present, the contrast, size distribution, and evolutionary timescales of granular convection are strongly modified [Title et al., 1992]. Convective flows also play a different character within the strong magnetic fields of sunspots (Figure 5). Recent high-resolution observations of sunspots have clarified a heretofore unsettled observational description of the fine-scale field structure and flows in sunspot penumbrae: Darker structures are regions of weaker, more horizontal magnetic fields that harbor most of the outward Evershed flow (see discussion in review by Lites [1997]). Nonetheless, we are still far from an understanding of these phenomena at a fundamental level where theory accurately describes the observed properties. Much more theoretical and modeling work is needed. Furthermore, the new observational magnetic field diagnostics have not yet supplied a comprehensive observational picture of the evolution of active regions at high angular resolution. That need will be met by planned ground- and space-based instrumentation (section 9).

1.3. Intense, Small-Scale Flux Tubes

Excluding sunspots, nearly all the magnetic flux visible in Figures 3 and 4 is concentrated into small (<200 km diameter), very intense (~1500 G) flux tubes (see review by Solanki [1993]). The small sizes, comparable to a density scale height in the solar photosphere, and strong fields, near equipartition of magnetic pressure...
Figure 4. The full-disk magnetogram at the upper left is representative of the data from which the synoptic chart of Figure 3 was constructed. The corresponding continuum image is shown at upper right. The magnetogram from the SOHO/Michaelson Doppler imager reveals the fine structure of the active region magnetic fields in the expanded view at lower left.
Magnetic Field Tracers
Active Region 8218, 13 May 1998

Figure 5. Monochromatic images at some wavelengths indicate the location (in the photosphere) and topology (above the photosphere) of the magnetic field without actual measurement of the field itself. These images (courtesy of T. Berger, Lockheed Martin Solar and Astrophysics Laboratory) represent very high angular resolution observations taken at the Swedish Solar Observatory, La Palma, Canary Islands, Spain, plus a simultaneous image in Fe IX-X emission from TRACE. Top row shows continuum intensity $I_c$ and G-band (430.5 nm, molecular CH, and other atomic species) filtergrams. Middle row shows Fe I 630.2-nm magnetogram and Ca II K-line (393.4 nm) filtergrams. Bottom row shows TRACE 171-Å image and La Palma Hα filtergram (656.3 nm). Brightenings in the G band (Figure 6) frequently reveal the location of intense photospheric flux tubes. The more diffuse chromospheric emission in the Ca II K-line also overlies the locations of magnetic flux. Extended chromospheric structures, sometimes tracing magnetic lines of force, are seen in Hα and in X rays such as the TRACE image. A compact flare is apparent in the lower left of the TRACE, Hα, and K-line images.
Figure 6. Enlarged images of the subarea outlined in Figure 5 show that brightenings in the G-band images reveal locations of small-scale photospheric magnetic fields. The G band is a series of absorption features around 430 nm which arise mainly from the CH molecule.

with the surrounding gas pressure, have long been inferred indirectly through involved analysis of the polarization of spectral absorption lines [Stenflo, 1973]. The highest-resolution images of the solar photosphere (e.g., Figure 6) reveal bright structures a few tenths of an arc second across or smaller (<200 km at the Sun; 1 arc sec is 725 km at the center of the solar disk as seen from Earth), thought to mark the location of individual flux tubes. Only very recently has it been possible to apply speckle image reconstruction techniques to polarization images in order to directly verify the tiny size of these flux concentrations (Figure 7).

Flux tubes are much more than an observational curiosity. Vigorous convective flows produce the granulation pattern of the photosphere visible at sizes 1000–2000 km in Figures 6 and 7. These flows buffet flux tubes that normally reside in the darker intergranular lanes. The dynamics of this interaction, believed to be the
Speckle Polarimetry in Quiet Network
Christoph Keller and Bruce Wilton (1997)

Figure 7. Speckle image reconstruction techniques have been applied to polarization imagery to produce magnetograms of extremely high angular resolution. These images show that photospheric flux tubes often reside in the dark lanes between granules (top image). Plots in the right column demonstrate that the features in both magnetic field and intensity are near the diffraction limit of the telescope (0.2" or about 150 km).

major source of heating of the outer layers of the Sun, has been the focus of considerable efforts in magneto-hydrodynamic (MHD) modeling of flux tubes [Steiner et al., 1998]. A variety of physical processes are likely candidates for transformation of the kinetic energy of photospheric convection into thermal energy of the upper layers, including the generation and dissipation of magnetoacoustic-gravity waves and multitudinous "nanoflares" occurring at very small scales which convert magnetic energy efficiently into heating of the outer layers of the solar atmosphere: the chromosphere and corona. (The chromosphere is a thin (~2000 km), highly dynamic layer between the photosphere and corona containing plasma that is somewhat hotter than the photo-
of magnetism in the quiet Sun is now apparent: The rate of flux emergence in the quiet Sun appears to be orders of magnitude greater than that from the total of active regions. This result is unexpected because when viewed in large scale, the evolution of solar magnetic flux is clearly dominated by active regions (Figure 3). Careful accounting of flux for emerging ephemeral active regions (small, emerging bipolar flux elements that develop and disperse away from active regions) has demonstrated an emergence rate at least 10 times that of active regions [Schrijver et al., 1997]. The emerging flux from ephemeral regions appears to sustain the assemblages of flux tubes at the network boundaries. At even smaller scales, much weaker flux elements emerge in greater numbers and frequency, such that their total emergence rate is at least a factor of 10 greater yet than that of ephemeral regions [Lites et al., 1996]. At present, it is unclear whether the weak internetwork fields are simply the result of convective shredding and mixing of the flux in the quiet network or if the vigorous convection near the solar surface produces its own small-scale dynamo [Schüssler, 1996]. The latter process would sustain these internetwork fields even in the absence of the deep-seated dynamo of the solar cycle. The attributes of the internetwork flux, intrinsically weak fields at small spatial scales, present a considerable challenge to observational technique. In order to learn more about them we need higher sensitivity to weak fields and higher angular resolution than are presently available. This need has spurred recent efforts for infrared solar polarimetry (section 7).

1.5. Remote Sensing of Solar Magnetism

The terrestrial influence by solar variability underscores the need to understand its magnetic origins at the Sun, with the eventual goal of predicting significant solar events and long-term variability. However, predictive numerical models of solar activity reminiscent of terrestrial climate and weather models are a very distant goal. Unlike the terrestrial atmosphere, the source of the solar luminosity and magnetic fields lies deep within the Sun. Magnetic flux must traverse the turbulent outer convective envelope of the Sun before escaping into the observable atmosphere. Although solar oscillations [Braun et al., 1998] and other observables provide some indirect observational information about the solar interior, the deterministic lower boundary condition of terrestrial atmospheric models is unavailable to us. Moreover, in the solar atmosphere a complex interplay occurs among convection, radiation, and magnetic fields in the presence of a strong gravitational stratification, leading to the richly diverse array of observed solar phenomena but also presenting us with a dauntingly complex physical system to decipher. Detailed, quantitative observations of the physical state of the solar atmosphere (temperatures, pressures, velocities, and magnetic fields) are the basic ingredients for furthering our understanding of these phenomena. Once a solar phenomenon is under-
Plate 2. The very quiet Sun imaged in the Ca II K-line (393.4 nm) in September 1996 is shown, along with nearly simultaneous, very high sensitivity magnetograms from the Advanced Stokes Polarimeter. The K-line image shows the chromospheric network, revealing the locations of magnetic fields in the photosphere below. A few patches of network flux are seen in the image at lower left. When this image is displayed so as to reveal the weakest flux elements (bottom middle image), small-scale, mixed-polarity internetwork flux is visible nearly everywhere. The magnetogram images are constructed from a scanned sequence of polarization spectra. One spectrum of Stokes $V$, corresponding to the highlighted slit position in the magnetograms, is shown at bottom right.
Plate 3. Soft X-ray images taken with the soft X-ray telescope (SXT) imager on the Yohkoh spacecraft show the growing loop of X-ray emission associated with a flare. Analysis of these and other temperature diagnostics [Tsuneta, 1996] indicate that magnetic reconnection is taking place above the X-ray loop seen here, and the bright region at the top of the loop is material cooling from temperatures in excess of 10 MK. Figure courtesy of M. Shimojo, the Yohkoh group of The Institute of Space and Astronautical Science, Japan, and NASA.
stood, one might then transfer this understanding to similar phenomena in other astronomical contexts which cannot be spatially and temporally resolved like their solar counterparts and which occur under conditions that cannot be reproduced in the laboratory.

Given the extraordinary, pervasive importance of magnetic fields in solar phenomena, this fundamental property of the solar plasma is perhaps the most important observable property of the solar atmosphere. Unfortunately, it is not yet possible to carry out in situ measurements of solar magnetic fields, so we must infer most of what we know about the solar atmosphere from the radiation it emits. (NASA plans the Solar Probe, which would approach within 4 $R_\odot$, sampling briefly the fields and plasma in the outer corona.) This review therefore addresses the remote sensing of solar magnetic fields. Capability for quantitative vector magnetic field measurements in the photosphere has dramatically improved in recent years, revealing the topology of fields as they emerge to form active regions and permitting reliable measures of their magnetic flux. Moreover, space-based measurements of the solar magnetic field have recently allowed long-term, continuous monitoring of the evolution of solar magnetic fields with unwavering image quality. These significant new observational tools are complemented by rapidly advancing theoretical and modeling capability for solar magnetohydrodynamics, and the combination of theory and observations has greatly widened our understanding of solar magnetism within the past decade.

This review emphasizes quantitative measurement of magnetic fields but begins in section 2 with a brief survey of intensity tracers of the field, some of which have long provided qualitative impressions of the field topology. Image sequences from the recent Solar and Heliospheric Observatory (SOHO) and Transition Region and Coronal Explorer (TRACE) space missions underscore the value of tracers of the field. Polarization measurements in spectral lines provide the basis for most quantitative measurements of solar magnetic fields. Section 3 briefly summarizes, from a classical standpoint, how magnetic fields create (and alter) polarization of spectral lines forming in the solar atmosphere. Section 4 reviews techniques for standard longitudinal magnetometry (i.e., those sensitive to the component of the field vector along the line of sight), which have been the mainstay of routine solar magnetic field measurements since the inception of the solar magnetograph by Babcock [1953]. Techniques for inferring the full field vector are reviewed in section 5, with emphasis on modern spectropolarimetry, which provides the most accurate estimates of all three components of the field vector. Also presented are a few specific attributes of the polarization measurement that largely determine the accuracy with which the vector magnetic field may be recovered from data.

Quantitative information about the magnetic field vector and the physical parameters that describe the thermodynamic state of the solar plasma are encoded into the amplitudes and shapes of the spectral lines. Robust techniques have been developed within the past decade to extract highly accurate measures of the magnetic field vector. The advent of high-quality polarimetric solar data has encouraged recent efforts to extract even more detail about the solar atmosphere from polarimetric data, i.e., gradients of the vector magnetic field and other parameters along the line of sight (LOS). These techniques are described in section 6. Polarization measurements in the infrared promise to reveal the nature of weak fields (section 7). Prospects for measurement of the field at heights well above the photosphere are given in section 8, and anticipation of future directions is presented in section 9.

2. TRACERS OF THE SOLAR MAGNETIC FIELD

We have already seen in Figure 6 that high-resolution imaging of the solar photosphere in the molecular G band reveals the locations of at least some of the intense photospheric flux tubes. This is one example of a tracer of the field that does not constitute an actual measurement of the field itself.

Some investigations of solar activity are greatly aided by observational information on the connectivity of magnetic fields above the solar surface. For example, changes in the connectivity of the magnetic field as a consequence of field line reconnection during a solar flare provide unique information about the flare process and its sudden release of magnetic energy. Under conditions prevalent in the solar atmosphere a magnetic field greatly suppresses thermal conduction (i.e., electron and ion transport) perpendicular to its direction, so that adjacent magnetic lines of force (i.e., magnetic "loops" in the solar atmosphere) may sustain very different thermal environments depending on the conditions along each loop, especially where the loop intersects the photosphere. This inhomogeneity perpendicular to magnetic lines of force causes magnetic loops to emit or absorb radiation differently from their neighbors, thereby enhancing the visibility of the field lines. A familiar analogy is the orientation of iron filings near the poles of a strong magnet: Some field lines are occupied by filings and others are not, leading to a convenient visualization of the magnetic topology. Visualization of field lines does not constitute a true measurement of the field, but actual field measurements above the photospheric level, especially for the full field vector, are problematic (section 8). Tracers thus represent a very important asset in the study of magnetic phenomena in the solar atmosphere.

Images alone may provide information on the magnetic connectivity. The X-ray images of the Sun shown in Plate 3 suggest that the connectivity of the field in the corona undergoes a substantial qualitative change as a result of a flare. Under most circumstances the corona is not opaque to X rays. This helps to reveal the full length
of magnetically confined structures but also hinders interpretation due to overlap of structures. Both optical depth and overlap effects are evident in the 171-Å X-ray image of Figure 5.

The magnetic topology in the corona is revealed only for those lines of force that contain plasma emitting at the observed wavelength. Most ultraviolet and X-ray emission lines become excited only at a specific, narrow range of temperature. Structures visible in the 171-Å band (8- and 9-times ionized iron, 0.9 MK) of Figure 2 may be distinct from those of the 5- to 20-Å band of the soft X-ray telescope on Yohkoh (Plate 4), typical of about 3 MK. Unlike the line emission that dominates UV and X-ray images, white-light coronal images (from either eclipse or coronagraphic observations) register continuum emission primarily from scattering by electrons. Hence white-light images such as those of Plate 1 are sensitive to the integrated coronal electron density along the LOS, but they are insensitive to temperature.

The density of large-scale structures in the corona plays an important role in determining the stability of coronal structures. Here the term large-scale refers to structures extending beyond a few tenths of a solar radius. Magnetic pressure dominates for smaller-scale structures near the base of the corona, but beyond one coronal scale height the plasma interacts significantly with the field to form the solar wind, i.e., to maintain field structures open to interplanetary space. Buoyancy effects may influence dynamic events such as the chromospheric white-light image seen in white light in Plate 1. Prior to a CME event, typical coronagraphic measurements reveal a bright helmet-streamer similar to that of the first image of Plate 1, the lower part of which has a dark bubble (the prominence cavity, also seen). A prominence will usually reside at the very bottom of the cavity. When destabilized, the whole system erupts rapidly outward, opening up the field lines of the helmet-streamer and sometimes ejecting prominence material with it out into the solar wind, as is seen as the expanding filamentary structure in the latter images of Plate 1. The key to the loss of stability of this system is probably the evacuated prominence cavity [Low and Hundhausen, 1995]. The cavity is characterized by reduced density and is surrounded by the denser material of the helmet-streamer, which weighs down on the cavity to counter the buoyancy arising from both its reduced density and higher field strength.

How does the cavity become partially evacuated? The field in the cavity likely has the form of a twisted magnetic "flux rope." The lines of magnetic force of the flux rope have a relatively direct topological linkage through the prominence below, but the rope is effectively isolated from both the hot coronal helmet-streamer above and the dynamic source of coronal heating from the photosphere below. The rope interior is thus likely to be thermally unstable. Material within the cavity at coronal temperatures may radiate, cool to chromospheric temperatures, and drain to the lowest reaches of the flux rope, forming a prominence and evacuating the cavity.

The large volume of the cavity allows it to act as a buoyant "balloon" that can support the prominence below. Once the helmet-streamer is no longer able to weigh upon the cavity, a CME eruption event will occur.

Intensity tracers of the field may be found in the lower atmosphere as well. Figure 5 presents a high-resolution photospheric continuum image of a sunspot in the low photosphere, in the chromospheric H α line, and in the longitudinal magnetic field. The continuum photospheric images show structures in the penumbra of the sunspot which indicate the general direction of the field at the photospheric surface. These structures end abruptly at the outer boundary, where it is believed that the magnetic field exits the photosphere but continues outward at higher levels of the atmosphere. The accompanying H α chromospheric image strongly suggests this outward continuation and also reveals complex structures, presumably indicating field orientation, which are not at all evident in the continuum image.

Optically thick structures such as those in Figure 5 may not always indicate the direction of magnetic lines of force. Because the scale height of plasma at photospheric and chromospheric temperatures is small (<200 km) relative to the length of the structures (>1000 km), plasma suspended in the corona by the field (i.e., a prominence, equivalently, a filament if viewed in H α in absorption against the solar disk; see Plate 5) likely reveals the location of local dips in the height of individual field lines, so that linear structures may even be aligned perpendicular to the field direction. Plate 4 presents a theoretical model of a prominence that shows just this behavior. Filaments often have "feet," localized extensions of prominence material extending downward toward the surface, which give the impression of a downward extension of field lines. Were this the case, the prominence plasma would rapidly drain from these structures because the cool plasma in the prominence has a scale height of only a few hundred kilometers. The more likely interpretation is that the feet reveal a system of concave upward field lines extending downward from the main body of the prominence, each of which has a concentration of dense plasma at chromospheric temperatures at its lowest point.

What magnetic topology stabilizes prominences for periods as long as months? This question is one of great current interest and debate. Whatever the actual field structure may assume, study of the H α tracers of a large sample of prominences has revealed a hemispheric handedness to the orientation of fine structures in filaments and their surroundings [Martin et al., 1994], indicating a global-scale organization of the magnetic field structure, likely an important indicator of the solar dynamo process. The fine structure of prominences illustrates the need for actual vector field measurements upward from the photosphere, which may then be considered along with tracers of the field at higher levels to build up a sound description of the field topology.
3. MAGNETIC FIELDS AND THE GENERATION/ALTERATION OF POLARIZATION

3.1. The Stokes Vector

Four quantities are needed to fully describe the state of polarization of electromagnetic radiation. The representation of polarized light by means of the Stokes 4-vector \( I = [I, Q, U, V] \) (the superscript \( T \) indicates matrix transposition) is a widely used basis. Conveniently, the components of \( I \) may be defined operationally in terms of intensities measured with ideal optical elements. Unlike representations of polarization based on the amplitudes of the electric vector (i.e., the complex Jones 2-vector), the Stokes vector also has the advantage that it describes partial polarization of radiation from multiple incoherent sources of light, the overwhelming majority of sources in the universe.

Presuming one has an ideal linear polarizer, ideal right- and left-handed circular polarizers, and an ideal detector, one may perform very simple measurements to arrive at each of the four elements of \( I \). The quantity \( I \) represents the intensity of light measured by a detector unaffected by polarization. With choice of a right-handed coordinate system having the +z axis in the direction of propagation, \( Q = I_x - I_y \), where \( I_x \) and \( I_y \) are the intensities measured with an ideal linear polarizer oriented along the x and y axes, respectively. If the coordinate system \((x', y', z)\) represents a rotation of the \((x, y, z)\) system about the z axis by \( \pi/4 \), such that \( x' \) falls halfway between \( x \) and \( y \), then one defines \( U \) as the measurement \( U = -I_x + I_y \). For measurements of light through ideal right- and left-handed circular polarizers designated as \( I_{rcirc} \) and \( I_{lcirc} \), respectively, Stokes \( V \) is defined as \( V = I_{rcirc} - I_{lcirc} \).

3.2. The Zeeman Effect

The Zeeman effect provides by far the most common and useful diagnostic of solar magnetic fields. The presence of a magnetic field in the solar atmosphere lifts the degeneracy of magnetic sublevels of the radiating atom, causing characteristic wavelength shifts (Zeeman splitting) of transitions among these sublevels. The orientation of the magnetic field also establishes a preferred reference frame for description of the interaction of the atomic states with polarization states of electromagnetic radiation. Because of the Zeeman splitting, the solar magnetic field alters the spectral shapes of emission or absorption lines and also imparts both circular and linear polarization. The detailed shapes of the spectral lines, in both intensity and polarization, emerging from the solar atmosphere depend upon atomic physics of the transitions being observed and upon the structure of the solar atmosphere. One must therefore formulate and solve the equation of radiative transfer for polarized radiation, including its interaction with the atomic system.

This theory may be approached from either quantum or classical viewpoints. A rigorous quantum mechanical formulation of this problem, based on the density-matrix formalism, was presented elegantly by Landi Degl’Innocenti and Landi Degl’Innocenti [1972]. It provides a consistent formalism of the problem of spectral line formation under conditions where the scattered photon is redistributed in frequency over the entire line profile (complete redistribution). A more pragmatic approach founded on an analogy of classical oscillators predated the quantum formulation and, for the most part, correctly describes the problem. A thorough development of this latter approach is also available [Jeffries et al., 1989]. The reader is referred to these authors, as well as to Stenflo [1994], for complete expositions of this problem. Presented below is a brief overview of the subject that may assist the reader with the material that follows.

It may be shown [Jeffries et al., 1989] that the Stokes vector satisfies a vector equation analogous to the scalar transfer equation of radiative transfer:

\[
\frac{dS}{dz} = -KS + j, \quad (1)
\]

where \( K \) and \( j \) are the \( 4 \times 4 \) absorption matrix and the emission 4-vector, respectively, replacing their corresponding scalar quantities in the transfer equation for unpolarized light, and \( z \) is the linear coordinate along the LOS increasing toward the observer. Several factors influence the form of \( K \): the magnetic field and, just as with unpolarized transfer, the physics of the atomic system giving rise to the spectral line; the state of motion (both macroscopic and microscopic) of the plasma; and the density of absorbers per unit volume. The emission term \( j \) may include the effects of scattering when it is important.

For most diagnostics of using the Zeeman effect for photospheric fields, the assumption of local thermodynamic equilibrium (LTE) is valid (i.e., collisional excitation, rather than scattering, dominates \( j \)), and the emission vector simplifies to

\[
J = K(B, 0, 0, 0)^T + B_s(K_{00}, K_{10}, K_{20}, K_{30})^T, \quad (2)
\]

where \( B_s \) is the Planck function. The problem of formation of photospheric lines influenced by the Zeeman effect then reduces to definition of the absorption matrix \( K \).

Consider the atomic system classically as a three-axis damped oscillator having a resonant frequency \( v_0 \) of the spectrum line. The oscillator may be represented equivalently by a system consisting of a linear oscillator along the direction of the magnetic field vector \( B \) (the \( \pi \) component) plus two circular oscillators of opposite directions, both lying in a plane perpendicular to \( B \) (the \( \sigma \) components), as shown in Plate 6. (This classical analogue provides a faithful representation of the quantum behavior only for atomic transitions where the upper and lower states have angular momentum quantum numbers \( J = 1 \) and 0, respectively. This configuration gives rise to the so-called normal Zeeman triplets.) This
is the preferred coordinate system because, owing to the action of the Lorentz force, the angular momentum of the orbiting electron undergoes a precession about $\mathbf{B}$ at the Larmor frequency $\nu_L = eB/4\pi mc$ ($e$ and $m$ are the electronic charge and mass, respectively). Incoming right-circularly polarized light excites a precession characterized by a magnetic moment parallel to $\mathbf{B}$, which represents the magnetic state of reduced energy $h(\nu_0 - \nu_L)$. Conversely, left-circularly polarized light is absorbed at the higher frequency ($\nu_0 + \nu_L$).

Plate 6 illustrates schematically the interaction of the oscillator with incident electromagnetic radiation. The longitudinal Zeeman effect occurs when a plane wave propagating in the direction of $\mathbf{B}$ excites both circular modes, $\sigma_R$ and $\sigma_L$, corresponding to right- and left-circular polarization, respectively, which are subsequently damped, leading to the absorption of circularly polarized light at $\nu_0 \mp \nu_L$. These circular modes correspond to the $\sigma$ components of the Zeeman pattern. In this special case of light propagating along $\mathbf{B}$ the incident electromagnetic wave has no oscillatory component along the linear oscillator parallel to $\mathbf{B}$ (the $\pi$ component), so absorption of that component cannot occur. Because the background continuum is unpolarized, the absorption of right- (left)-circular polarization results in an excess of left- (right)-circular polarization in the residual intensity as evidenced by negative (positive) Stokes $V$ in Plate 6. It must be remembered that the Stokes spectra are usually plotted on a wavelength scale instead of a frequency scale used in Plate 6.

The other special case is $\mathbf{B}$ transverse to the direction of propagation, the transverse Zeeman effect. The propagating wave excites both $\pi$ and $\sigma$ components. The former is unshifted in wavelength by the magnetic field, and the latter only influences the state of linear polarization of the incident beam that is perpendicular to $\mathbf{B}$, or along $y$. The three components of the Stokes $Q$ profile (up-down minus left-right linear polarization) are identified with the linear polarizations resulting from the absorption process.

The shapes of the Stokes profiles of solar spectral lines contain information about the full vector $\mathbf{B}$: both the longitudinal (along the LOS) and transverse (perpendicular to the LOS) components. Note that $I$ is invariant to a rotation of the transverse component of the field by $\pi$ rad. The Zeeman effect is insensitive to this ambiguity in the field orientation (the "180° azimuth ambiguity"), so that other physical input to the problem is required (section 6). Note also that when the Zeeman components overlap somewhat, the magnetized plasma becomes birefringent. Near the centers of opaque absorption lines, Faraday rotation, or the magneto-optical effect, may affect $I$.

Figure 8 presents Stokes spectra of a narrow wavelength region containing two transitions arising from Zeeman-sensitive neutral iron (Fe I) in the solar atmosphere and two transitions due to O$_2$ in the Earth's atmosphere. The spectral profiles of the Fe I lines reveal a fundamental property of polarization spectra as dictated by the analogy of Plate 6: They are nearly symmetric about $\nu_0$ in $I$, $Q$, and $U$ and are antisymmetric in $V$. One spatial position near the umbra-penumbra boundary is highlighted in both the sunspot image and the spectra. Actual shapes of the observed Stokes profiles at this location are given by the dots in Figure 9. At the center of the Stokes $V$ profile of the 630.25-nm line is a small reversal of opposite sign. This feature arises from the magneto-optical effect.

### 3.3. The Hanle Effect

The Hanle effect (see Stenflo [1994] for a full exposition in the context of solar magnetic fields) describes the interaction of polarized radiation and atoms under conditions well removed from the regime of the Zeeman effect. For this reason it has proven to be a useful diagnostic of weak magnetic fields high above the photosphere and in weak-field regions of the photosphere itself. Relative to light scattered by an identical atom in a field-free region, the Hanle effect introduces both a rotation of the plane of polarization and (usually) a reduction of the net polarization of the scattered light.

Three criteria must be fulfilled in order for the Hanle effect to be operative. (1) The Larmor frequency must satisfy $\nu_L \sim A_{ij}$, where $A_{ij}$ is the spontaneous emission
rate of the atomic transition. For permitted transitions, $|B|$ must be in the range of a few to a few tens of gauss. (2) The plasma densities must be low enough and the radiation intensity must be high enough that radiative excitation of the atom is more common than collisional excitation that depolarizes the emitted radiation. (3) The radiation field incident upon the scattering atoms must be anisotropic. The latter two criteria are usually satisfied at heights a few hundred kilometers (or greater) above the solar surface. These criteria severely restrict the realm of applicability of the Hanle effect.

In the quantum description of the Hanle effect the incident anisotropic radiation causes the populations of the excited magnetic states to differ. Because the field is weak, the wave functions describing each state overlap in energy, and there may be interference among them. The interaction of the quantum system representing each magnetic state differs with respect to the polarization of the incoming radiation; hence when quantum interference is present, the scattered radiation represents a mixture of polarization states.

Like the Zeeman effect, a classical oscillator analogy is also possible for the Hanle effect in the case of the normal Zeeman triplet, as depicted in Plate 7. Plate 7a represents a scattering system illuminated anisotropically by a plane wave propagating along the $x$ axis. Assume the magnetic field is directed along the $y$ axis and the observer is positioned along the $z$ axis. For pure scattering in the absence of a magnetic field the observer would measure pure linear polarization oriented parallel to the $y$ axis: No excitation and reemission of polarization occurs along $x$ because the incident wave has no corresponding component of the electric or magnetic field. As with the Zeeman effect, we represent the oscillator by a linear oscillator along $y$ and two oppositely rotating circular oscillators in the $x$-$z$ plane. The magnetic field causes a mixing of the field-aligned state with the circular states, causing some reemission polarized along $x$. Therefore, in this geometry, the Hanle effect reduces the net polarization of the scattered radiation but does not alter its orientation along the $y$ axis. Depolarization alone also occurs if the magnetic field is oriented along the $x$ axis (Plate 7b): In this case the magnetic field mixes the radiatively excited circular oscillator states into the linear state aligned with the magnetic field. The more interesting special case is where the field is aligned along the LOS, i.e., along $z$ (Plate 7c). The observed light results from the circular states only. A pure scattering process in the limit of $B = 0$ results in observed linear polarization along $y$ because both circular states are excited in phase, with the resultant polarization perpendicular to the LOS. The magnetic field alters the phase between these oscillators, and the finite lifetime of the reemission process causes the damping of both circular states. This time-dependent variation of the oscillators results in both a reduction of the net polarization and a rotation of the plane of polarization.

For solar applications of the Hanle effect, often the anisotropy arises principally from limb-darkening of the solar disk, because the spectral lines employed form in the upper photosphere and the chromosphere where the solar disk effectively occupies 2$\pi$ sr. For Hanle studies of prominences well above the limb the smaller apparent size of the Sun at the scatterer is also a major source of anisotropy. In either case, the net anisotropy is small, resulting in net linear polarization of the scattered radiation, typically a few percent or less. The Hanle effect has proven to be a very useful diagnostic of solar prominences (see section 8) because they meet all three criteria cited above.

### 3.4. Coronal Forbidden Lines

The spectrum of the inner corona at visible wavelengths reveals a number of emission lines [Billings, 1966], nearly all of which arise from magnetic dipole transitions in the ground terms of multiply ionized atomic species. Such transitions have small spontaneous emission rates, of the order of 100 s$^{-1}$, whereas the Larmor frequency is $1.4 \times 10^6 |B|$. The magnetic states are thus well separated in frequency, and there is no quantum interference as in the Hanle effect. Scattering of anisotropic radiation by a magnetic dipole produces linear polarization that is either aligned with the projection of the magnetic field vector on the plane of the sky or perpendicular to it. This is a geometrical effect only, as the degree of polarization is essentially independent of the strength of the field. The geometry of Plate 7c produces no net polarization. If $\theta$ is the angle between the direction of the incident radiation and the field, then the radiation scattered by a magnetic dipole is linearly polarized along the field direction for $\theta < \cos^{-1} \sqrt{1/3}$ (the Van Vleck angle [see House, 1974]) or perpendicular to the field direction for larger angles.

Like the Hanle effect, forbidden line polarization is reduced if collisions provide excitation of the emitting atoms comparable to or greater than the radiative excitation. Unlike the Hanle effect, the degree of polarization (and the discontinuous change in orientation due to the Van Vleck effect) are independent of the $|B|$. Nonetheless, the observed orientation of the plane of forbidden line polarization can be a useful measure of the orientation of the field. It may also be possible to mea-

---

**Plate 5.** (opposite) Prominences are islands of relatively cool (7000 K), dense plasma suspended by magnetic fields in the hot corona. The larger, red image, which is at top when the plate is correctly oriented, shows a prominence in emission above the solar limb as seen in the H$\alpha$ line. The smaller sequence of images at bottom shows H$\alpha$ images of a prominence rotating past the solar limb. Prominences seen in absorption on the disk are called filaments.
Plate 6. The classical analogue of the atomic system as a system of orthogonal oscillators serves to illustrate the essentials of the Zeeman effect: the interaction of atoms with polarized radiation. The case depicted has optically thin absorption with the field aligned along the $x$ axis, so that Stokes $U = 0$, and hence is not shown.

Plate 7. The classical oscillator analogy is used to illustrate the alteration of polarization due to mixing of atomic states in the Hanle effect.
Figure 8. Polarization spectra of a sunspot measured by the Advanced Stokes Polarimeter [Elmore et al., 1992] are shown. Spectral observations of the Stokes parameters $I$, $Q$, $U$, and $V$ characterizing the complete state of polarization of a narrow range of a sunspot spectrum near 630 nm are shown in the lower four panels. The top right panel shows the location of the spectrograph slit over the sunspot. The two broader spectrum lines due to absorption by neutral iron atoms form in the solar atmosphere and show both solar structure and polarization. The two narrow, polarization-free lines arise from oxygen molecules in the Earth's atmosphere. The horizontal lines near the top and bottom are present for calibration purposes. The white lines indicate the location of the spectral profiles shown in Figure 9.
Figure 9. Polarization profiles and least squares fits. A set of individual Stokes spectra from the location indicated in Figure 8 are shown (dots), along with least squares fits to the two solar spectrum lines (solid curves). The wavelength scale is indicated in nanometers. The extracted magnetic field parameters for this location in the sunspot are \(|B| = 2206 \, G\), \(\phi = 140^\circ\), and \(\psi = 134^\circ\). At this spatial location the optical thickness of the spectrum lines is considerably larger than unity, so that Faraday rotation, also known as the magneto-optical effect, causes the direction of linear polarization to rotate as a function of wavelength near the centers of spectral lines. It also causes an anomalous "reversal" near the center of the Stokes \(V\) profile at lower left. This effect is accounted for in the analysis, as the solid line of the least squares fit matches the observed reversal of \(V\). The magneto-optical effect actually helps to establish the optical depth of the lines in the least squares analysis.

Sure \(|B|\) via the small degree of circular polarization due to the Zeeman effect (section 8).

4. LONGITUDINAL SOLAR MAGNETOMETRY

The classical analogue of the Zeeman effect (Plate 6) presented in the previous section illustrates how circular polarization is produced by the component of the magnetic field along the LOS. The term in the absorption matrix \(K\) in (1) that converts Stokes \(V\) from Stokes \(I\) involves \(v_b \cos \psi\), where \(v_b = v_L/\Delta \nu_D\), \(\psi\) is the angle between \(B\) and the LOS, and \(\Delta \nu_D\) is the Doppler broadening width. This angular dependence of Stokes \(V\) on \(\cos \psi\) is the basis for longitudinal magnetometry.

In the visible spectrum, solar magnetic fields usually produce a much larger Stokes \(V\) signal than that of the linear polarization \(Q, U\) (Figure 8). The weaker the field, the smaller the ratios \(Q/V\) and \(U/V\) are. Indeed, for a normal Zeeman triplet in the limit of weak fields or small \(v_b\) it may be shown [Jefferies et al., 1989, equations 45 and 47]

\[
V = v_b \cos \psi \frac{\partial I}{\partial v}, \quad Q, U \propto (v_b \sin \psi)^2, \quad (3)
\]

where the displacement from line center \(v_0\) is \(v = (v - v_0)/\Delta \nu_D\). In contrast to the quadratic dependence of \(Q, U\) on \(|B|\), the linear dependence of Stokes \(V\) on \(|B|\) underlies the historical preference for longitudinal magnetometry over full vector magnetometry: The Stokes \(V\) signals are simply much larger. For most lines encoun-
tered in the visible spectrum the approximation leading to (3) is valid for fields outside of the umbræae of sunspots.

Equation (3) shows that the Stokes V signal is proportional to the frequency derivative of the intensity profile. Hence V maximizes in the line "wings" (i.e., away from line center) where the slope of the Stokes I profile is greatest and leads to the antisymmetric V spectral profiles seen in Figures 8 and 9. In essence, longitudinal magnetographs use optical devices that modulate only the circular polarization and then detect this modulated signal in the wings of solar absorption lines. For example, a quarter-wave retarder (the "modulator") may be rotated continuously in the beam to convert circularly polarized light into a harmonic variation of linear polarization. When the modulator is followed by a linear polarizer (the "analyzer"), the modulated polarization is converted to an easily detected harmonic variation of intensity. Routine acquisition of solar magnetograms of "weak" solar magnetic fields outside of sunspots began with the Babcock [1953] magnetograph about a half century ago. Longitudinal magnetograms have been recorded daily at several observatories spanning several solar cycles. The Kitt Peak magnetograms [Livingston et al., 1976; Jones et al., 1992], represented by the full disk image in Figure 4 and the synoptic map of Figure 3, comprise one widely used data set of particularly high quality and long duration.

Babcock's original magnetograph used a spectrograph to isolate the very narrow range of wavelengths of Zeeman-sensitive solar absorption lines. The lineage of Kitt Peak magnetographs has also used spectrographic dispersion, and the current system [Jones et al., 1992] images the entire spectral line to improve on the accuracy of the estimates of magnetic flux. Rapid modulation and detection of the circular polarization are needed to minimize spurious polarization signals from the unavoidable rapid fluctuations in image distortions due to the Earth's atmosphere (atmospheric seeing). Historically, imaging detectors were not fast enough to minimize the effects of seeing. The spectrographic observations permitted high-precision polarimetry at the expense of spatial coverage of the solar disk. Advances in detector and data processing technology greatly improved the spatial resolution, but the spectrographic slit must still be scanned over the solar image, and this limits the time resolution of the magnetographic observations, typically 1 hour for the entire Sun in the current Kitt Peak system. Studies of evolution in the magnetic field are compromised, especially those studies relating to small-scale structure. Video imaging technology; image restoration ("destretching" the seeing distortions of the image due to seeing in post facto data analysis); and greatly improved narrowband, wavelength tunable filter technology allowed imaging longitudinal magnetographs to fulfill this need [Wang et al., 1998; Title et al., 1992]. Imaging technology has been somewhat easier to implement in space than spectrographic observations, and the absence of seeing relaxes somewhat the requirements for rapid modulation. (At the high angular resolution possible from space, the evolution of the solar surface itself determines the modulation rate needed to achieve a given polarimetric precision. The precision of polarimetry from the Michelson Doppler imager (MDI) on SOHO (SOHO/MDI) is so limited.) The cost of imaging magnetometry is, of course, the limited spectral information available. It is primarily this lack of spectral information that prevents most longitudinal magnetographs, and imaging vector magnetographs, from providing quantitative measures of actual field strength.

Simultaneous circular polarization in two spectral lines differing principally in their sensitivity to the Zeeman effect revealed [Stenflo, 1973] at once that Babcock's weak fields outside of sunspots were instead quite intense (~1500 G), and thus the magnetized plasma is isolated spatially to very small concentrations within much larger, nearly field-free zones. If this were not the case, the spectral shapes of the polarization profiles and their relative polarization would be inconsistent with the small observed degree of polarization. With this observational result, the concept of flux tubes occupying only a fraction of the spatial resolution element (small magnetic filling factor) was born. The key to this analysis was the simultaneous recording of two spectral lines. Precise spectropolarimetric observations of many lines simultaneously at low angular resolution, though at very high spectral resolution, with the visible light Fourier transform spectrometer (FTS) at Kitt Peak further validated the flux tube concept and provided information on the temperature structure of the magnetic elements (see review by Solanki [1993]).

Limitations of standard longitudinal magnetographic observations are revealed in the foregoing story of the discovery of flux tubes. Without angular resolution adequate to completely resolve spatially the individual magnetic elements, the observed degree of circular polarization is diluted by the essentially unpolarized radiation from the nonmagnetized component within the field of view. The inferred "field strength" involves the observed Stokes I profile (equation (3)), which does not faithfully represent the magnetized region alone. Thus low magnetic filling factors translate into erroneous, low inferred magnetic field strengths. Even today, many magnetographic observations are misrepresented in terms of |B| (gauss), when in fact they represent some measure of magnetic flux within the resolution element. To make matters worse, ever-present scattering in the optical system and time-variable seeing further reduce the magnetic filling factor. The final caveat concerns the calibration of longitudinal magnetographs. Usually, (3) is assumed. It breaks down in the strong fields of sunspots: In the limit of complete Zeeman splitting the degree of circular polarization is independent of |B|. Furthermore, many magnetographs measure polarization and intensity at only one or two wavelengths, with no point-by-point information about ∂I/∂θ. A calibration using spatially
TABLE 1. Criteria for Precision Vector Field Measurements

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Value</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polarimetric precision</td>
<td>$&lt;10^{-3}$ spectral continuum</td>
<td>avoid systematic errors from instrumental sources</td>
</tr>
<tr>
<td>Signal-to-noise ratio (SNR)</td>
<td>$&gt;10^6$</td>
<td>attain SNR in Stokes $Q$, $U$, $V &gt; 10$ in weak flux regions outside of sunspots</td>
</tr>
<tr>
<td>Spectral resolution $\lambda/\Delta\lambda$</td>
<td>$&gt;2 \times 10^5$</td>
<td>spectrally resolve subtleties of Stokes profile shapes</td>
</tr>
<tr>
<td>Spectral coverage</td>
<td>$\geq 2$ lines simultaneously</td>
<td>isolate effects of thermodynamics from those of magnetic field</td>
</tr>
<tr>
<td>Temporal cadence</td>
<td>video rate for spectral imaging (ground-based)</td>
<td>atmospheric seeing has power to $\sim 500$ Hz</td>
</tr>
<tr>
<td>Angular resolution</td>
<td>$&lt;1^&quot;$ (700 km on Sun)</td>
<td>magnetic structures exist on range of scales down to at least 100 km; 1&quot; is useful for study of intermediate-scale structure of active regions</td>
</tr>
</tbody>
</table>

averaged profiles is adopted, but this cannot account for large fluctuations in the shape and depth of solar absorption lines. For example, note the dramatic weakening of the Stokes $I$ profiles precisely at those locations of flux tubes (as indicated by Stokes $V$) at the lower portion of the spectrum in Figure 8. The weakening of the lines is due to the enhanced temperatures within flux tubes. Several modern instruments acquire and process data in ways that account for the calibration difficulties. The latest Kitt Peak magnetograph [Jones et al., 1992] measures resolved spectral profiles and processes them in real time, and the SOHO/MDI instrument [Scherrer et al., 1995] measures 5 wavelengths throughout the line. Neither of these enhancements, however, compensates for the very important issue of magnetic filling factor. To date, no solar observations have fully resolved the fine-scale structure of the magnetic field, so the reader should not only be extremely wary of field strengths quoted for magnetographic measurements, but he or she should also be skeptical of quoted sensitivity limits. In general, only involved analysis of spectrally resolved Stokes profiles (section 5) will reveal the true $|B|$ in spatially unresolved or partially resolved observations.

In spite of these limitations, and in spite of the absence of sensitivity to the transverse magnetic field, longitudinal magnetometry remains an extraordinarily valuable observational tool for solar physics. Both the long-term synoptic measurements discussed in section 1 (Figure 3) and the time series of filter magnetograms from SOHO/MDI [Schrijver et al., 1997] illustrate the importance and utility of longitudinal magnetometry.

5. VECTOR MAGNETOMETRY

Even if one were able to measure the longitudinal component of the solar magnetic field with high precision, one would still lack crucial information about the structure of the field hidden in its two other components. For numerous problems of solar physics these components are all-important. One example of considerable interest currently is the evolution of the twist, or shear, in the magnetic field as an active region emerges and spreads in to the solar corona. This process has the potential not only to reveal the nature of the fields below the surface, perhaps even carrying information about the dynamo deep within the Sun that produced them, but also to reveal much about the interaction of these magnetic flux ropes with the preexisting coronal fields that produce solar flares and coronal mass ejections. For this and many other problems, clearly it is far preferable to work with measures of the full vector magnetic field, even if that information is available only near the solar surface.

Unlike longitudinal solar magnetometry, vector magnetometry implicitly demands quantitative measures of the three components of the magnetic field. A substantial error in any component compromises the vector measurement. Polarimetry must have precision sufficient to extract each component of the vector to the accuracy needed. Thus "qualitative vector field measurement" is somewhat of an oxymoron.

Fortunately, the capability for precision remote sensing of the solar vector field has been realized within this decade. The new observations have stimulated interest in the theory and modeling of solar MHD processes and have spurred developments for greater sophistication in analysis of polarimetric data (section 6). The ingredients for this advance are embodied in the data shown in Figure 8 and are summarized in Table 1.

The advent of fast, low-noise charge-coupled device (CCD) detectors, along with digital video processing hardware, permitted rapid acquisition of full polarimetric observations of resolved spectral profiles as shown in Figure 8. From the ground, fast sampling of the modulated polarization signal is essential in order to suppress spurious "cross talk" from Stokes $I$ into $Q$, $U$, and $V$ arising from the rapid fluctuations of seeing. Kilohertz or greater sampling rates are needed to entirely suppress seeing cross talk to measure very small polarizations (of the order of $10^{-4}$ to $10^{-5}$), as achieved by the unique Zurich imaging polarimeter (ZIMPOL) device [Povel et al., 1994]. The ZIMPOL system permits exploration of the "second solar spectrum" [Stenflos, 1997]: the multitude of weakly polarized spectral features that appear in linear polarization of the spectrum near the edge of the
Figure 10. This scatterplot compares the results of detailed fitting of the Stokes spectral profiles (abscissae) with results from a more simplified analysis (ordinates) when full spectral resolution is not available. Each point represents one spatial location where the Stokes spectrum is measured. Departures from the straight lines indicate errors in the magnetograph analysis. Inferred field strengths are compared on the left, and inclinations to the line of sight (LOS) are compared on the right. The magnetograph analysis severely underestimates the field strength, leading to inferred fields that are much more highly inclined to the LOS than is actually the case.

Solar disk, arising from scattering of the anisotropic radiation field (see section 3). However, for photospheric lines of the visible solar spectrum, the polarimetric precision of $10^{-3} I_c$ listed in Table 1 is adequate for many studies of magnetic fields. This is achieved with conventional imaging at video rates, aided by simultaneous measurement of both of the modulated polarization states [Lites, 1987; Skumanich et al., 1997]. Precision spectropolarimetry has not yet been achieved in space, where the absence of atmospheric seeing relaxes the required data rate (Table 1). Efforts are under way to project this capability into space (see section 9). In the meantime, the potential of ground-based vector magnetometry has hardly been tapped.

Vector magnetographs based on narrowband imaging of $I$ at one or perhaps two wavelengths in the wing of a Zeeman-sensitive absorption line have been operated for some time prior to the advent of precision spectropolarimetry. Why go to the trouble of measuring full spectral profiles? Figure 10 [Lites et al., 1994] compares the measured $|B|$ and inclination angle to the LOS, $\psi$, from detailed analysis of spectral profiles (abscissae) with a commonly used magnetographic analysis as applied to a 0.01-nm range of wavelengths in the line wing of these same data (ordinates). (Methods for analysis of full Stokes spectral data provide error estimates for the magnetic field (section 6) that have been verified both by numerical tests [Skumanich et al., 1992; Westendorp Plaza et al., 1998] and by analysis of real solar data [Lites et al., 1998]. For typical spectropolarimetry as analyzed to produce Plate 4, errors in components of $B$ are typically a few percent, depending on the signal-to-noise ratio (SNR) of the $Q$, $U$, and $V$ spectra.) The magnetographic analysis encounters substantial errors in both $|B|$ and $\psi$, as anticipated from the discussion at the end of the previous section. Unlike the magnetographic analysis, the detailed profile analysis is able to estimate magnetic filling factors and to account for the point-to-point variations in thermodynamics. Neglecting these issues leads to severe underestimates of $|B|$, which in turn lead to measures of $\psi$ which are much more inclined to the LOS than is actually the case. These errors notwithstanding, the field azimuth, the orientation angle of the magnetic field in the plane of the sky, is recovered to reasonable accuracy by the magnetographic analysis. This orientation angle has been used as a diagnostic for changes in the vector field [Wang et al., 1994]. However, vector magnetic fields are most commonly desired in the local solar frame defined by the local direction of gravity. Errors in $\psi$ propagate into errors in the field azimuth measured in the local solar frame.

Plate 4 illustrates how the interplay between theory and observed solar vector magnetic fields has considerably broadened our understanding of emerging magnetic flux and the evolving structure of the overlying atmosphere [Lites et al., 1995]. The small emerging active region indicated at the left in Plate 4 was followed during its rotation across the solar disk by the advanced Stokes polarimeter [Elmore et al., 1992] providing vector field measurements, by the soft X-ray imager on the Yohkoh spacecraft, and by ground-based imaging instrumentation at the Swedish Observatory in the Canary Islands. Because the region was small and rather simple, the comprehensive body of observational fact draws one compellingly toward a novel interpretation of the evolution of the small delta sunspot region at left (a delta
sunspot is one which includes umbrae of opposing magnetic polarity within a single penumbral region), an interpretation with far-reaching consequences. The delta region is not emerging as a simple, bipolar nest of magnetic loops. Instead, it is a highly twisted, perhaps even nearly closed magnetic structure. Above the delta spot, the Lorentz force associated with the strong currents of this structure are able to thermally isolate cool (7000 K), dense plasma (prominences) from the surrounding 10^6 K coronal plasma and support this material against gravity. This interpretation may not only explain much of the behavior of delta sunspots which are responsible for many of the most intense solar flares, but it also may provide a useful paradigm for the behavior and fine structure of prominences and for the evolution of large-scale magnetic field evolution leading to CMEs.

The vector field measurements of the photosphere shown at the left in Plate 4 are the key that unlocked this paradigm. Precise polarimetry allows one to locate the true magnetic inversion line separating the opposing polarities within the delta sunspot, even when the region was far removed from the center of the solar disk. The inversion line is indicated by red in the visible light perspective image and by a white line on the planes displaying the vector magnetic field. An inversion line is always present under a prominence. At the time of these observations a small prominence was developing over the inversion line bisecting the delta sunspot as seen in the H alpha image from the Canary Islands. The polarity and inclination of the vector field are indicated by the color bar: Yellow hues represent field vectors emerging from the surface, and blue hues indicate returning vectors. Arrows in the upper plane show the orientation of the vectors, as does the "hair" in the lower magnetic field image. Note in the delta region that the field vectors diverge away from the inversion line, rather than converging toward it, as would be expected from a dipolar magnetic field that does not deviate greatly from a potential field. At the inversion line the field is nearly horizontal to the surface and parallel to the inversion line.

The theoretical magnetostatic model [Lites et al., 1995; Lites and Low, 1997] visualized at the lower right in Plate 4 embraces the most important features of the observations. The model magnetic field is toroidal and closed. A few inner lines of force are highlighted as blue tubes: They traverse from positive (yellow at the base) to negative (blue) polarity along rather simple but twisted paths in the volume above the plane. The single outer white tube highlights one of a class of field lines that define this model: It winds several times around the blue field lines and passes through several local minima of its height above the plane. Highlighted in red are the local minima of a large collection of field lines of the class indicated in white. They occur above the inversion line at the plane and possess many features observed in real prominences. Indeed, this model reproduces the "inverse" polarity of prominence fields: The projection of the horizontal component of the field onto the direction perpendicular to the axis of the prominence is opposite that defined by the potential field. The potential field is uniquely defined by the observed vertical component of the field at the photosphere. The Hanle effect has been applied to polarization measurements of prominences [Leroy et al., 1984; Bommier et al., 1994] (see section 8), showing that the great majority of prominences have inverse polarity fields.

6. ANALYSIS TECHNIQUES FOR SOLAR POLARIMETRY

A wealth of information about the solar atmosphere is contained in precision spectropolarimetric data such as those of Figure 8. The instrumentation required to record such complete Stokes spectral data is now available and is becoming more commonplace, so this review concentrates on the analysis of full spectropolarimetry rather than on older methods for dealing with more limited data, which involve restrictive approximations.

The most effective analysis techniques involve least squares fitting of the Stokes spectra using solutions of (1) for models of the solar atmosphere along the LOS [Del Toro Iniesta and Ruiz Cobo, 1996]. These models must not only describe the magnetic field vector, but they must also faithfully mimic the thermal and dynamical structure of the atmosphere as it varies from point to point on the Sun. This is accomplished by representation of the solar atmosphere by models comprising a few parameters that capture, to some degree of approximation, the variation of the real solar atmosphere both spatially and along the LOS. One fairly simple model, the so-called Milne-Eddington (ME) atmosphere, admits an analytic solution to (1) [Rachkovsky, 1962; see also Landolfi and Landi Degl'Innocenti, 1982], thus allowing the computationally efficient "inversion" [Stumanich and Lites, 1987] of tens of thousands of Stokes spectra typically contained in a map like that shown in Figure 8. The ME atmosphere assumes a linear variation of the source function B, with optical depth \( \tau \), and other parameters (B, line broadening, line strength, LOS velocity) are held constant. ME fits to synthetic Stokes spectra as generated from fully variable model atmospheres demonstrate that the ME analytic solution yields quite accurate estimates of the LOS means of B [Stumanich et al., 1992; Westendorp Plaza et al., 1998]. An example of a fit to one set of Stokes spectra is given by the solid lines in Figure 9. These spectra correspond to the spatial position highlighted in Figure 8.

Current instrumentation rarely, if ever, allows one to resolve magnetic structures in the solar photosphere, but subresolution structure must be accounted for in the analysis of Stokes spectra. It is especially important in active regions outside of sunspots, where unresolved flux tubes comprise the bulk of the magnetic flux. A first-order approximation commonly used in analyses of spectropolarimetric data is to assume a two-component atmosphere: One component is magnetized, and the other
is not. In the simplest construct the unmagnetized component is assumed to produce Stokes I spectra identical to that of the mean quiet Sun. The least squares procedure allows an admixture of the unpolarized Stokes I spectrum, usually derived from a relatively field-free region of the spatial map, to the fit to the Stokes profiles. This combination of magnetized and unmagnetized spectra mimics the presence of several unresolved but identical and homogeneous magnetic structures and/or the presence of an unpolarized stray light spectrum. Thus the least squares fitting of the Stokes spectra may yield a first-order estimate of the magnetic filling factor. Estimates of the filling factor are possible only because of the subtle inconsistency of the observed Stokes I spectral profile with the shapes and amplitudes of polarization profiles of Q, U, and V. However, without this correction for filling factor/stray light, serious errors in the vector field will be assured if one demands a fit to the full Stokes vector.

The experience of least squares fitting of many millions of Stokes profiles, for literally hundreds of active regions, suggests that the ME model, as modified by the filling factor correction described above, produces an excellent representation of true solar vector fields in the overwhelming majority of cases. Confidence in the verity of the inversion results derives from a number of the properties of the data, as enumerated by Lites et al. [1994]. Both the continuity of the derived vector field from one observed spatial location to the next and the invariance of the nearly vertically oriented plage fields even when observed near the limb of the Sun [Martinez Pillet et al., 1997] are strong indicators of a valid analysis, within the limitations of the observations and the ME atmospheric model.

Least squares fitting procedures also allow robust estimates of the uncertainty of each of the parameters being fit. Such error estimates immediately flag spectra for which the routine fails to find a satisfactory fit. For Stokes spectra with SNR > 1000, failure is rare, provided one supplies a good initial guess for the parameters of the fit. Without a good initial guess, least squares fitting procedures may settle on a secondary minimum in the typically 10-dimensional parameter space. Genetic algorithms [Charbonneau, 1995] provide one means to isolate the locality of the global minimum in parameter space. Although the genetic algorithm is computationally too inefficient for the full fit, it provides an excellent guess for the standard least squares procedure.

The genetic algorithm has proven most valuable in refinement of the analysis of spatially unresolved Stokes spectra. Very occasionally, observed Stokes polarization profiles Q, U, and V depart dramatically from the highly symmetric spectral shapes typically found (e.g., Figure 9), such that the Stokes V profile, usually antisymmetric about line center, even takes on the appearance of a wavelength-shifted symmetric Q or U profile. In some cases, these abnormal profiles may be explained by the unresolved, yet physically distinct, structure of the magnetic field. For example, these abnormal profiles may occasionally be found in sunspots viewed away from the center of the disk in the vicinity of the sign reversal of the line-of-sight component of the field vector in the penumbra. It is now well known [Lites, 1997] that penumbral harbor fine structure in which the fields differ in strength and inclination to the surface, and also in their bulk motion. Extending the ME model to allow two fully variable magnetic components plus an unmagnetized one within the observed pixel will often allow one to find a satisfactory fit to the observed Stokes profiles [Skumanich et al., 1996; Bernasconi and Solanki, 1996]. However, the expansion of the number of free parameters of the fit greatly increases the occurrence of secondary minima in the parameter space. Application of the genetic algorithm to the initial guess has been found to greatly ease the problem of secondary minima in this case [Skumanich et al., 1996].

The ME analysis does not admit LOS variations in the magnetic field, in the line-of-sight velocity component v_{LOS}, or in the other important thermodynamic properties. Variations in v_{LOS} are responsible for departures from symmetry of Stokes Q and U and from antisymmetry of V about line center: Note the asymmetries present in the observed spectra in Figure 9. High-precision, low-noise spectropolarimetry contains useful information about variation along the LOS of all the parameters that characterize the solar atmosphere. If this information can be extracted from observations, it will be an extremely important aid to deciphering the physical system. For example, a knowledge of local gradients of all components of B allows one to extract the electric current density ($J = (c/4\pi)\nabla \times B$), an important diagnostic of the solar plasma. Furthermore, it provides a glimpse of the three-dimensional structure of the atmosphere, albeit limited in extent to the few scale heights in the atmosphere over which the diagnostic spectral lines form.

Advances in techniques for interpretation of Stokes spectra have followed in step with recent advances in instrumentation. New techniques for fitting the spectral profiles embrace more realistic atmospheric models, allowing free variation of conditions that influence the emergent I. A number of groups are actively engaged in refinement of these sophisticated inversion methods. For example, one technique, the Stokes inversion based on response functions (SIR) [Ruiz Cobo and Del Toro Iniesta, 1992], has been applied to sunspot observations, the results of which are shown in Plate 8 [Westendorp Plaza et al., 1997, 1998]. At the lowest heights sampled by the analysis of this positive polarity sunspot (the bottom panel of Plate 8), horizontal or even negative polarity patches of flux are found near the outer edge of the sunspot penumbra. These patches also harbor downward flowing material. The more standard analysis of sunspot Stokes spectra, which does not account for variations along the LOS, senses only higher levels of the atmosphere, represented by the top and middle panels in
Plate 8. Detailed analysis of the shapes of the Stokes polarization spectral profiles yields information on gradients in the vector magnetic field. This image (courtesy of C. Westendorp Plaza and J. C. del Toro Iniesta) indicates the inclination of the magnetic field to the local vertical direction of a sunspot at three heights in the solar atmosphere. Note that near the outer edge of the sunspot in the bottom panel, patches of opposite-polarity flux exist which are not present in the upper two panels (see text.)
Plate 8, where the field at the outer edge of the penumbra has a positive vertical component. This newly discovered three-dimensional structure of sunspot magnetic fields may be the long-sought sink for the outward flow from sunspots, the Evershed flow. It is revealed only through detailed analysis of subtleties in the shapes of the Stokes spectra.

A more radical approach to the analysis of Stokes spectra has been proposed: the microstructured magnetic atmosphere (MISMA) [Sánchez Almeida et al., 1996, 1997]. This scenario is able to reproduce the large asymmetries observed in Stokes profiles in a variety of solar features without invoking the rather extreme (sometimes unphysical) LOS variations [Sánchez Almeida and Lites, 1992] required for the same degree of asymmetry in a more homogeneous atmosphere. In the MISMA scenario the atmosphere is highly structured on scales much smaller than the mean free path of a photon: For example, it might be a field-free background threaded by many very small magnetic flux tubes, with the flux tubes harboring a significant flow velocity. The MISMA hypothesis is appealing, especially for interpretation of flux in the quiet Sun, where recent theoretical models [Cattaneo, 1999] suggest the field may be highly tangled on very small scales.

The Zeeman effect carries no information on the parity of the component of the vector field transverse to the line of sight \( B_z \). That is, \( B_z \) is said to be ambiguous by \( \pi \) rad because the linear polarization introduced by the Zeeman effect responds identically to either \( B_z \) or \(-B_z\). Resolution of this ambiguity in the field azimuth angle \( \phi \) requires additional information about the field or its origin. Such information may take the form of an assumption: The azimuth may be chosen such that it is closest to that of the potential field defined uniquely by the locally vertical component of the field \( B_z \) [Sakurai et al., 1985; Cuperman et al., 1992], or one may choose the azimuth such that discontinuities of the field vector are minimized [Lites et al., 1995]. It is possible to extract the part of the vertical electric current that is independent of the ambiguity [Semel and Skumanich, 1998]. This current aids in evaluating the validity of a particular resolution of the ambiguity and in isolating true discontinuities in electric currents that may exist in the atmosphere. Additional observational information may also be added to help resolve the ambiguity. At chromospheric levels, scattering polarization may help [Landi Degl’Innocenti and Bonnier, 1993]. If simultaneous observations are available at two heights in the atmosphere, the condition \( \mathbf{V} \cdot \mathbf{B} = 0 \) leads to a unique solution [Wu and Ai, 1990]. In some cases the field connectivity defined by tracers of the field at and above the chromosphere (section 2) resolves the ambiguity for some photospheric fields. At present, there seems to be no magic bullet. In large measure, the resolution of the azimuth ambiguity remains an art, and a combination of methods is used to resolve the ambiguity for any given spatial map of the vector field. Thus errors of \( \pi \) may persist in the transverse field component of vector magnetograms.

7. INFRARED SPECTROPOLARIMETRY

The energy perturbation of atomic states is proportional to the strength of the magnetic field \( |B| \). The splitting in wavelength is thus proportional to \( |B| \lambda^2 \), but the Doppler broadening of an atomic transition is proportional only to \( \lambda \). As a result, the Zeeman splitting relative to the Doppler broadening of spectrum lines varies as \( |B|/\lambda \). Spectrum lines forming in intrinsically weak solar magnetic fields of less than 1000 G do not show a separation of the Zeeman components at visible wavelengths, but the splitting of a similar line in the infrared can be nearly complete (i.e., can be comparable to or larger than the Doppler broadening; see Figure 11). Infrared measurements offer clear advantages for the study of weak field regions in the solar atmosphere. When the magnetic splitting is so large that the individu-
ual Zeeman components are separated in wavelength as in Figure 11, the degree of polarization is maximal and independent of field strength. In the large splitting limit the degree of linear polarization is also comparable to that of circular polarization, so infrared observations relax the stringent limits on polarimetric precision imposed upon vector magnetometry in the visible spectrum. If the splittings are large, it is often possible to clearly distinguish multiple magnetic components within the resolution element by their differing Zeeman splittings [Rüedi et al., 1992]. The task of extracting the properties of differing field components may be addressed with less model dependency than for the visible wavelength observations cited in section 6. Additionally, atmospheric seeing decreases with wavelength, as do the polarizing properties of optical reflections. For these reasons, there is now a great interest in instrumentation for solar polarimetry in the infrared [Kuhn and Penn, 1995].

There is a common misconception regarding polarimetry at visible wavelengths: It is often stated that quantitative vector field measurements are not possible when the Zeeman splitting satisfies $v_b < 1$, i.e., when $|B| < 600$ G for typical Zeeman diagnostic spectral lines. This is a significant underrating of visible light diagnostics. Recent spectropolarimeters routinely achieve SNR $\approx 10^3$. At this level, residual random noise, not intrinsic solar conditions, limits the capacity for visible light measurement of weak fields. Realistic simulations [Westendorp Plaza et al., 1998] show that spatially resolved weak-field vectors may be recovered for field strengths as small as 100 G. Of course, the errors in the measured $|B|$ increase with decreasing field strength, with increasing noise, or with decreasing filling factor. As long as the weak fields remain unresolved spatially, skepticism will justifiably persist as to the validity of the analysis of polarimetry in the visible.

Until high-resolution visible-light spectropolarimetry is available in space (section 9), infrared diagnostics provide the least ambiguous information about physical conditions in the weak internetwork fields [Lin, 1995a, 1999]. These observations in the neutral iron lines at 1.56 µm forming low in the photosphere strongly suggest that the internetwork fields are intrinsically weak, a few hundred gauss, and unlike the intense fields in flux tubes, histograms of the occurrence of fields of a given strength are rather broad (Figure 12) [Lin, 1995b]. These infrared observations mainly describe the fields within the more prominent internetwork flux elements. The strength of the nearly ubiquitous weaker, mixed-polarity internetwork flux still remains largely unknown. High-resolution, high polarimetric precision measurements from space from Solar-B (see section 9) will constrain the sizes of these elements and may very well provide information on intrinsic field strengths. One must await infrared polarimetry at much higher angular resolution to determine distributions of field strengths for the small-scale, weak internetwork patches that one sees as
a “salt-and-pepper” pattern in very sensitive magnetograms [see, e.g., Lites et al., 1999]. Analysis of the Hanle depolarization of certain spectral lines near the limb of the Sun suggests that much weaker fields (∼100 G) are pervasive in the quiet photosphere [Faurobert-Scholl, 1993]. It remains to be seen if any part of the known salt-and-pepper internetwork fields gives rise to the Hanle depolarization.

The field strength drops rapidly with height above the photosphere due to the strong gravitational stratification. If suitable infrared diagnostic spectrum lines of chromospheric and coronal fields can be found, one may be able to measure the intrinsically weak fields there. The neutral magnesium lines at 12 μm provide very sensitive field strength measurements [e.g., Hewagama et al., 1993] in the upper photosphere [Carlsson et al., 1992]. The near-infrared He I lines (1.083 μm) provide important optically thin chromospheric diagnostics, but their Zeeman splittings are comparable to those of chromospheric lines in the visible. These lines are intrinsically broad due to the low atomic weight of helium, and absorption is weak. Furthermore, both the He I and 12-μm lines are not excited everywhere on the solar disk, so that their usefulness is limited to a few classes of solar phenomena. In summary, no infrared lines have yet been found that provide adequate diagnostics for weak fields in the upper solar atmosphere.

Because many of the prominent science questions of solar physics are driving the observations toward higher angular resolution, a major hurdle for infrared solar magnetometry will be to overcome limitations on resolution dictated by diffraction limits of telescopes, which vary as λ. Very large infrared telescopes are needed to resolve the small-scale structure of the solar atmosphere. This limitation is of concern particularly for measurements in the thermal infrared. Larger-aperture infrared telescopes are also desirable because of the lower-energy flux and lower quantum efficiency of infrared detectors. The size, efficiency, and availability of infrared array detectors have increased steadily in recent years, making infrared vector magnetometry progressively more useful.

8. MAGNETIC FIELD MEASUREMENTS ABOVE THE PHOTOSPHERE

Most solar magnetograms sample magnetic fields very close to the visible solar surface. This limited picture of the field is insensitive to some of the most interesting physics of solar magnetic fields. Even slightly above the photosphere the pressure exerted by the magnetic field may greatly exceed the gas pressure (B = 8πPgas/B^2 ∼ 1), resulting in magnetic domination of atmospheric structure and giving rise to dynamic phenomena such as solar flares and CMEs. Simultaneous measurements of B at two heights would help to resolve the ambiguity in field azimuth (section 6). They also would facilitate efforts to extrapolate fields to higher levels. The Lorentz force usually dominates other dynamical forces when B ≪ 1, so that currents tend to flow along the direction of the magnetic field. When this is strictly true, the field is said to be force-free, and extrapolations are much more tractable for force-free fields [Amari et al., 1997]. Thus it is of great scientific interest to pursue analogous diagnostics of the chromospheric and coronal magnetic fields.

Substantial challenges confront efforts to measure magnetic fields above the photosphere. From the perspective of ground-based observation, there are only a handful of spectrum lines that form significantly above the photosphere, and these lines are not as sensitive to the Zeeman effect as those lines commonly used as photospheric field diagnostics. Owing to the larger non-thermal velocities, the chromospheric lines are intrinsically broader than photospheric lines and thus produce smaller polarization signals. Most important, though, outside of sunspots the magnetic fields of the chromosphere and corona are commonly very much weaker than photospheric fields, owing to the expansion of concentrated fields of flux tubes with height. Of course, this expansion eases the problem of accounting for small magnetic filling factors, but regions of rather uniform chromospheric or coronal magnetic field can have highly variable emission. Finally, in the chromosphere, most diagnostic lines are optically thick, are strongly influenced by scattering (non-LTE effects), and form over extended distances along the LOS. These issues complicate the analysis of chromospheric Stokes spectra.

Although there is as yet no fully satisfactory means of measuring vector magnetic fields above the photosphere, routine longitudinal magnetogram measurements are carried out daily for the Ca II line at 854.2 nm forming in the upper photosphere and low chromosphere (Figure 13). These measurements graphically illustrate that the field has “fanned out” considerably from the concentrated fields of the photosphere, mostly because of the combination of highly localized fields at the photospheric level and the rapidly decreasing density with height in the atmosphere. Similar diagnostics of the longitudinal field are provided by the Balmer sequence line Hβ [Zhang, 1994]. The neutral magnesium lines at 518 nm are sensitive to vector fields in the upper photosphere. Analysis of Mg I Stokes spectra using the simple ME approach has proven to be problematic [Lites et al., 1988], but more detailed treatment of the atmospheric model as discussed in section 6 will likely improve inversions of chromospheric Stokes spectra [Soares-Nava et al., 1998].

For chromospheric lines forming in the visible and near ultraviolet, small Zeeman splittings necessitate substantially higher polarization sensitivity than corresponding photospheric measurements. Infrared measurements would be ideal for this purpose, but few suitable spectral lines have yet been identified (section 7).
Prominence magnetic fields represent an interesting and important science target. Conditions in prominences are such that the Hanle effect has been applied very successfully to infer the vector magnetic field, revealing the dominance of inverse polarity prominence fields (section 5). Analysis of prominence fields via the Hanle effect are complicated by (1) the model dependency of the analysis (the density dictates the degree of collisional depolarization that must occur in addition to the depolarization from the Hanle effect); (2) the fine-scale structure of the prominence, which is usually not resolved in the plane of the sky and is not resolved along the LOS for prominences observed above the limb; (3) the degree of polarization usually being small; and (4) the azimuth ambiguity needing to be resolved. The analysis of two lines observed simultaneously [Bommier et al., 1994] allows one to determine all three components of the field vector and, in many instances, resolve the azimuth ambiguity. Most Hanle effect analyses so far have been directed at prominences above the solar limb, where often there is considerable foreshortening of the prominence structures due to their predominantly east-west orientation. A new analysis of filaments observed in absorption against the disk in the He I 1083-nm line [Lin et al., 1998] suggests that the Hanle effect might prove a useful diagnostic for the fine-scale structure of $B$ in prominences and for tracking its evolution for the full rotation across the disk.

Because of the intrinsic weakness of coronal fields and the difficulty of observing the weak coronal emission outside of eclipse, there are few prospects for coronal field measurements. This situation may soon change in light of the recent identification of forbidden coronal emission lines in the infrared with high Zeeman sensitivity [Kuhn et al., 1999]. It is anticipated that the circular polarization in these lines will allow one to measure longitudinal magnetic fields in the low corona over active regions where field strengths are highest. Very high polarimetric sensitivity will be required, which is difficult to achieve given the low brightness of the corona. Furthermore, the relatively large anisotropy of the incident radiation field at coronal heights may require a more detailed non-LTE treatment for interpretation of the observed circular polarization [Casini and Judge, 1999]. In combination with measurement of the linear polarization, which in the strong field limit (i.e., coronal forbidden lines) indicates directly the orientation of the field in the plane of the sky (section 3), a powerful observational tool for coronal physics may be forthcoming. As with any diagnostic of the corona, one must contend with the problem of variation of the atmosphere (including $B$) along the LOS, and this will ultimately limit utility of coronal field diagnostics off the solar disk.

The microwave continuum provides another diagnostic of magnetic fields in the low corona. The diagnostic utilizes the rapid variation of the opacity near the gyroresonance frequency ($\nu_{gy} = eB/2\pi m_e c$) for the precession of the magnetic moment of an electron in a magnetic field. The microwave brightness temperature falls rapidly with increasing frequency above a limiting frequency (or some multiple thereof). The limiting frequency is directly proportional to the field strength. This technique can yield an indication of those temperatures of the upper atmosphere at which specific field strengths are encountered, but so far it seems to be limited to the very strong fields encountered over sunspots [Lee et al., 1993].
TABLE 2. Solar Magnetic Fields: New Observational Capabilities

<table>
<thead>
<tr>
<th>Observational Concern</th>
<th>Science Problems</th>
<th>Planned Instrumentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-resolution, precision</td>
<td>Physics of flux tubes; chromospheric/coronal heating; weak fields in quiet Sun</td>
<td>Solar-B (joint Japan/U.S. space mission); Solar-Lite (U.S./German initiative for very high angular resolution)</td>
</tr>
<tr>
<td>polarimetry</td>
<td></td>
<td>Solar-B; SOLIS (ground-based full-disk spectropolarimetry); SOHO/MDI (exists)</td>
</tr>
<tr>
<td>Long-term, uninterrupted precision</td>
<td>Active region evolution; twist/shear of fields; indicators of dynamo</td>
<td>Advanced Solar Telescope (AST) (large aperture ground-based telescope with coronagraphic capability)</td>
</tr>
<tr>
<td>vector field measurements</td>
<td>weak internetwork flux; coronal vector fields</td>
<td>Trace (exists); Solar-B; STEREO</td>
</tr>
<tr>
<td>High angular resolution; infrared</td>
<td>indicators of field above photosphere; heating mechanisms</td>
<td></td>
</tr>
<tr>
<td>polarimetry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EUV/X-ray tracers of field topology,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>evolution</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initiative</th>
<th>Web Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOLIS</td>
<td><a href="http://www.nso.noao.edu/solis">http://www.nso.noao.edu/solis</a></td>
</tr>
<tr>
<td>AST</td>
<td><a href="http://www.nso.noao.edu">http://www.nso.noao.edu</a></td>
</tr>
<tr>
<td>SOHO</td>
<td><a href="http://sohowww.nascom.nasa.gov">http://sohowww.nascom.nasa.gov</a></td>
</tr>
<tr>
<td>TRACE</td>
<td><a href="http://vestige.lmsal.com/TRACE">http://vestige.lmsal.com/TRACE</a></td>
</tr>
<tr>
<td>SMI</td>
<td><a href="http://www.hao.ucar.edu/smi">http://www.hao.ucar.edu/smi</a></td>
</tr>
</tbody>
</table>

9. FUTURE DIRECTIONS FOR INVESTIGATION OF SOLAR MAGNETIC FIELDS

Table 2 summarizes some major instrumentation initiatives for solar magnetic field measurement, the observational needs that they address, and the relevant science goals as indicated in this review.

A theme running through all of solar physics is the need for high angular resolution. The magnetic field in the solar photosphere is structured on very small scales, and many processes of importance to heating and dynamics of the upper solar atmosphere occur principally at these small scales. It is therefore of utmost importance to describe the dynamical behavior of magnetic fields at these scales, accompanied by observations of the consequences in the upper layers, also at high angular resolution. The upcoming Japanese/U.S. mission Solar-B will provide continuous, high-resolution imaging polarimetry and spectroscopic polarimetry, along with imaging and spectroscopy of the upper layers. When this mission is launched in the 2004–2005 time frame, it will provide the first space-based high-resolution quantitative vector field measurements. Efforts are also under way to develop space instrumentation of even higher resolution, which will actually resolve the tiny, intense photospheric flux elements and their accompanying dynamical interaction with the solar convective motions (Solar Lite).

In the meantime, the resolution of field measurements from the ground continues to improve as a result of advances in off-line image processing, particularly the speckle imaging techniques discussed above and phase diversity imaging [Lähdahl and Scharmer, 1994], which uses simultaneous images (in focus and slightly out of focus) to estimate the seeing-induced distortions of the wave front. However, the resolution of precision vector field measurements based on spectropolarimetry will improve significantly only with real-time image correction: adaptive optical systems. Progress in solar adaptive optics has been steady, but a functioning adaptive optical correction still appears to be some years in the future. The scientific benefit of proposed large-aperture solar telescopes (e.g., Advanced Solar Telescope (AST)) will depend crucially on the availability of a well-functioning adaptive optical system. Indirect methods which permit one to infer certain aspects of the small-scale flux elements without actually resolving them will continue to be an important source of information about the solar magnetic field.

Another theme in the study of solar magnetic fields is the growing symbiosis among observation, analysis, and theory. Only within the last few years have new concepts regarding the fundamental nature of solar magnetic fields begun to link processes and phenomena observed on large and small scales. This unifying theme underpins the Solar Magnetism Initiative (SMI): a community initiative for support of a broad-based science program regarding the origin and nature of solar magnetic fields. SMI will support some ground-based instrumentation, but a major component will support efforts to improve interpretation techniques for observations from all available sources and to foster further theoretical and modeling efforts.

The long-term synoptic measurements of the Sun have proven to be of immense scientific value. The recovery of the SOHO spacecraft renews hope that the uniform-quality, uninterrupted sequence of magneto-
grams from that extraordinary space mission will continue to be available. The National Solar Observatory has embarked on development of instrumentation to provide frequent vector magnetograms of the entire solar disk, based on quantitative spectropolarimetric measurements rather than on the more qualitative data represented in Figure 3. This instrumentation is part of the Synoptic Optical Long-Term Investigations of the Sun (SOLIS) package, which should be completed in about 2001. The SMI program proposes to expand the SOLIS full-disk vector magnetic field measurements to a network of three stations at widely separated geographic longitudes in order to maximize the temporal coverage of the observations.

Any information one may gather about magnetic fields in the interior of the Sun will be enormously helpful in understanding the dynamo process and the interaction of the fields with the solar convective envelope. Helioseismology, the study of the oscillatory modes of the Sun itself, has provided precise information about the interior structure of this star. Local helioseismology has also begun to provide some indication of the subsurface structure in active regions [e.g., Braun et al., 1998], but this information is limited in its diagnostic content: The oscillations are influenced not only by magnetic fields, but also by the associated thermodynamic perturbations. The major contribution of helioseismology to the understanding of fields in the solar interior may derive from statistical variations of the gross properties of the interior with the solar cycle.

Finally, the magnificent new images of the corona provided by SOHO and TRACE demonstrate that tracers of the field topology down to scales of about 1000 km in the corona are finally available on a continuous basis. The diagnostic content of these data is still largely untapped, and these missions promise to continue to provide images for years to come. A major challenge to solar physics is now to combine them with precision vector field measurements at the photosphere in order to understand the evolution of fields upward from the solar surface to the outer corona.

Acknowledgments. A number of colleagues were enormously helpful in supplying data for illustrations and the illustrations themselves: Joan Burkepile, Tom Berger, Christoph Keller, David Elmore, Jack Harvey, Karen Harvey, Alice Lecinski, Haosheng Lin, B. C. Low, Jesper Schou, Dick Shine, Andy Skumanich, and Andy Stanger. I also thank Paul Charbonneau, Andy Skumanich, Alexandra Tritschler, Dick White, and others for providing references and suggestions and Roberto Casini for carefully reading the manuscript. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Jim Horwitz was the Editor responsible for this paper. He thanks Karel Schrijver and one anonymous reviewer for the technical reviews and Jean Braun for the cross-disciplinary review.

References


Baker, D. N., J. H. Allen, S. G. Kanekal, and G. D. Reeves, Disturbed space environment may have been related to power satellite failure, Eos Trans. AGU, 79(40), 477, 482–483, 1998.


---

B. W. Lites, National Center for Atmospheric Research, 3450 Mitchell Lane, Boulder, CO 80301. (lites@ucar.edu)