Astron. Astrophys. 277, 639–647 (1993)

ASTRONOMY AND ASTROPHYSICS

Infrared lines as probes of solar magnetic features

VI. The thermal-magnetic relation and Wilson depression of a simple sunspot

S.K. Solanki[1], **U. Walther**[1], **and W. Livingston**[2]

[1] Institute of Astronomy, ETH-Zentrum, CH-8092 Zürich, Switzerland
[2] National Solar Observatory, NOAO *, P.O.Box 26732, Tucson, AZ 85726, USA

Received December 3, 1992; accepted April 30, 1993

Abstract. We analyse the relationship between the temperature and the magnetic vector, as derived from 1.56 μm spectra, in a simple sunspot. Due to the high Zeeman sensitivity of the g = 3, 1.56 μm line, we can study this relationship throughout the sunspot. We confirm the field-strength, B, vs. temperature, T, relationship found by Kopp & Rabin (1992). In addition, we also find a linear relation between the magnetic inclination angle, γ', and T. An analysis based on the assumption of magnetohydrostatic force balance gives an estimate of the Wilson depression, Z_W, throughout the sunspot. Our analysis supports a jump in Z_W of 200–500 km at the umbral boundary and agrees with a relatively constant Z_W in the penumbra. In addition, we constrain various gradients of magnetic parameters and judge the strength of magnetic curvature forces. For example, we set tight limits on the vertical gradient of γ', the magnetic inclination angle to the vertical, in the penumbra. Finally, we discuss the consequences for, e.g., the model proposed by Sánchez Almeida & Lites (1992) to explain the broad-band circular polarization measured in sunspots.

Key words: sunspots – Sun (the): magnetic fields — MHD

1. Introduction

The radial temperature distribution of a sunspot in magnetohydrostatic (MHS) equilibrium is related to the structure of its magnetic field. Such a relationship was first predicted by Alfvén (1943) and has been theoretically and observationally investigated by, e.g., Cowling (1957), Dicke (1970), Maltby (1977), Martínez Pillet & Vázquez (1990, 1993) and Kopp & Rabin (1992). See Martínez Pillet & Vázquez (1993) for a detailed overview of the observational work on this subject. If magnetic curvature forces are neglected, then theory predicts a simple relationship between the vertical magnetic component, B_z, or alternatively the field strength, B, and temperature, T, at a given height. Observations in the visible have so far been restricted to the umbra, where they indeed support a linear relation between the squared field strength, B^2, and T. Martínez Pillet & Vázquez (1990) pointed out that this relationship can be used to estimate the Wilson depression, Z_W, in the umbra. Recently, the relationship between B and T has been investigated using spectra at 1.56 μm (Livingston 1991; Kopp & Rabin 1992). In the umbra these spectra confirm the results obtained in the visible. Due to the superior Zeeman sensitivity of the 1.56 μm lines (see, e.g., Solanki et al. 1992a, Paper II of the present series), Kopp & Rabin (1992) were able to extend the empirical B^2 vs. T relationship out to the visible boundary of the sunspot. Unexpectedly, outside the umbra, they saw strong departures from the predicted linear relationship, and they only analysed the umbral portion of their data in detail.

In the present paper we also investigate the connection between the magnetic field and temperature of a sunspot, as derived from 1.56 μm spectra. We improve on previous investigations in three respects:

1. We use measurements of both B and the inclination angle, γ', of the field to the vertical,[1] so that we can investigate the relationship between the magnetic vector and T.
2. We analyse in detail the B_z vs. T relationship individually at each observed position in the sunspot, including the penumbra.
3. We consistently obtain relevant physical parameters from a set of theoretical model atmospheres, thus significantly reducing the simplifications and assumptions introduced into the analysis.

Send offprint requests to: S.K. Solanki

* Operated by the Association of Universities for Research in Astronomy, Inc. (AURA) under cooperative agreement with the National Science Foundation.

[1] We write γ' instead of γ in order to distinguish it from the angle between the magnetic vector and the line of sight generally denoted by γ (cf. Paper V, i.e. Solanki et al. 1992b).

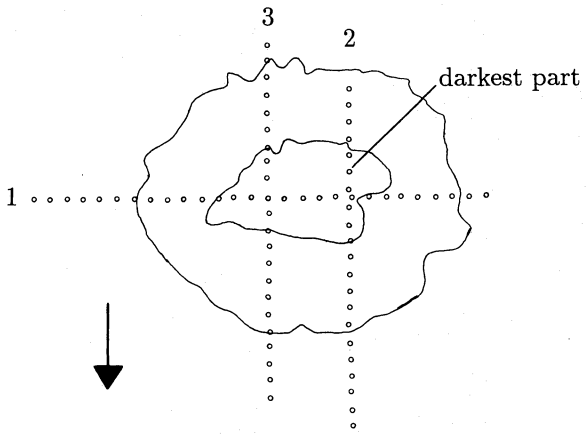

Fig. 1. Contours of the umbral and penumbral boundary of the sunspot. The small circles represent the positions at which spectra were obtained. The three slices through the sunspot are numbered near their starting positions. The arrow points towards the centre of the solar disc

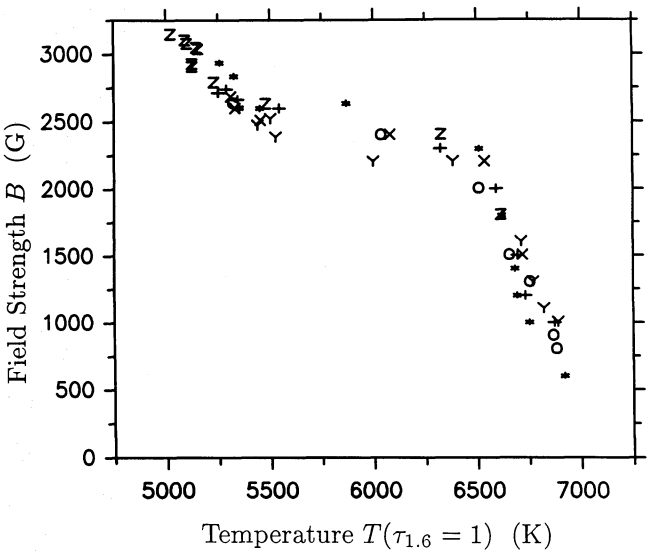

Fig. 2. Magnetic field strength, B, vs. temperature at unit continuum optical depth, $T(\tau_{1.6} = 1)$. The symbols represent different halves of the 3 sunspot slices. Stars = left (first) half of slice 1, plusses = right (second) half of slice 1, 'Z' = limbward half of slice 2, 'Y' = diskward half of slice 2, 'X' = limbward half of slice 3, open circles = diskward half of slice 3

Using the $g = 3$ line at 1.56 μm has the additional advantage that it measures the field strength close to the continuum-formation level, at which T is determined. Due to the generally small vertical gradient of B in sunspots, the B measured with the 1.5648 μm line is a very good approximation of B ($\tau_{1.6} = 1$). Thus, one of the assumptions underlying the interpretation of the B–T relationship, namely that B and T refer to the same height, is fulfilled to a high degree. Note that all τ values in the present paper refer to solar disc centre. Like previous investigators we average over magnetic and thermal fine structure.

2. Observed relationships and stray light

2.1. Observed relationships

The current investigation is based on data described and analysed in Paper V of the present series (Solanki et al. 1992b). The data are composed of Stokes I and V spectra of the Landé $g = 3$ Fe I line at 1.5648 μm, obtained in a relatively symmetric, mature sunspot close to solar disc centre ($\mu = \cos\theta = 0.985$). Outlines of the sunspot and of its umbra are shown in Fig. 1. Three slices were taken through the spot (marked 1,2,3 in Fig. 1). Each circle marks the position at which a spectrum was obtained. For the current analysis we only consider the data points lying within the visible contours of the sunspot. In addition to these, for a part of the analysis we have also used six spectra obtained in the immediate vicinity of the point marked 'darkest part'. As described in Paper V, B and γ' were determined by fitting the observed profiles with calculated I and V profiles using an inversion technique. For the present investigation we only need the B, γ' and relative continuum intensity values at each spatial position.

We first convert the continuum intensity, I_c, into temperature, T, using the Planck function, i.e. by solving the equation

$$\frac{I_c}{I_{c0}} = \frac{e^{hc/\lambda k_B T_0} - 1}{e^{hc/\lambda k_B T} - 1} \qquad (1)$$

for T. Here h is Planck's constant, c is the speed of light, k_B is Boltzmann's constant and λ is the wavelength at which I_c is measured. T_0 and I_{c0} are quiet sun values. In the remainder of the paper subscript 0 refers to the quiet sun. The I_c used for this purpose is already corrected for the non-magnetic stray light determined in Paper V. In accordance with the Eddington-Barbier relation $T_0 = T_0(\tau_{1.6} = 1)$, which we take from the quiet sun model of Maltby et al. (1986), so that the derived T should correspond closely to $T(\tau_{1.6} = 1)$ at the relevant position in the sunspot. Due to the opacity minimum at 1.6 μm, $\tau_{1.6} = 1$ refers to deeper and hotter layers than $\tau_{0.5} = 1$. The Eddington-Barbier relation is, of course, only an approximation and an additional source of uncertainty.

Figure 2 shows the resulting B vs. T relationship. It is very similar to the one seen by Kopp & Rabin (1992) for six sunspots. In particular, our data also show a plateau at the umbral boundary. The presence of such a plateau simply reflects the fact that the temperature changes rapidly at the umbral boundary, while the field strength does not. The different symbols represent the six half-slices through the sunspot. They suggest that although the relationship is relatively independent of azimuthal position in the spot, some systematic trends nevertheless exist. For example, compare the stars with the 'Y' symbols. The largest scatter in the figure is seen at the umbral boundary and at the outer edge of the penumbra. Since these positions are the most susceptible to stray light, it is not clear whether the scatter is intrinsic to the sunspot, or is due to stray light (see Sect. 2.2).

In contrast to previous investigators, we also have γ' values at our disposal. Remarkably, these show a linear dependence on T (Fig. 3). The umbral boundary is only identified by the increased scatter and the lower density of points. Apparently, the

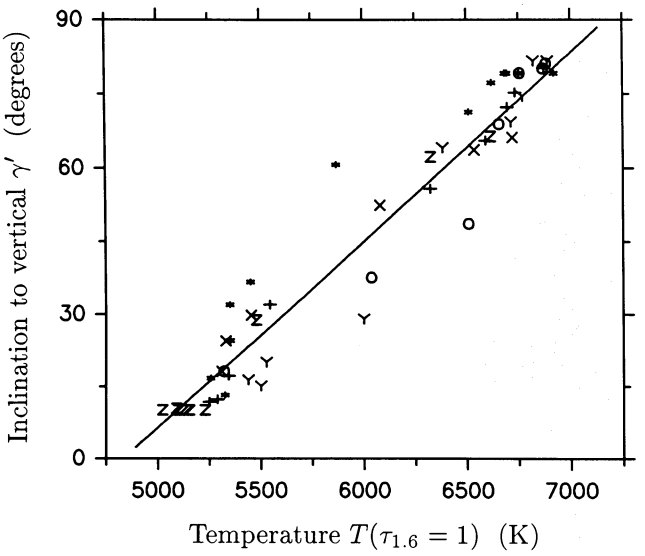

Fig. 3. Angle of inclination, γ', of the magnetic vector to the surface normal vs. $T(\tau_{1.6} = 1)$. The symbols have the same meaning as in Fig. 2. The straight line is a least-squares fit to the data

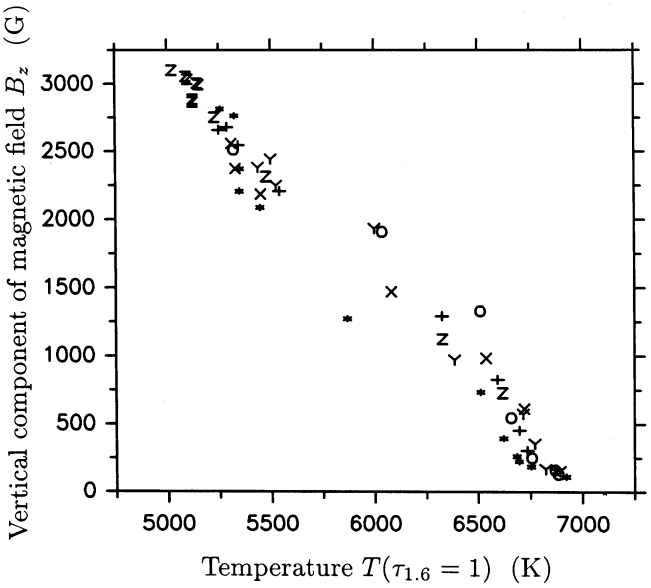

Fig. 4. Vertical magnetic component, B_z, vs. $T(\tau_{1.6} = 1)$

field changes considerably in inclination at the umbral boundary over a short horizontal distance. If the field were close to potential, then γ', like B, would form a plateau at the umbral boundary. The linear relation between γ' and T greatly influences the B_z vs. T and B_r vs. T diagrams (Figs. 4 and 5, respectively), where B_z is the vertical and B_r the radial component of the field (following the arguments given in Paper V, we consider an untwisted field). Basically, Figs. 2–5 give the following picture. Although the field strength does not change when passing from the umbra to the penumbra, the magnetic orientation becomes rapidly more horizontal (at the, possibly variable, height of line formation).

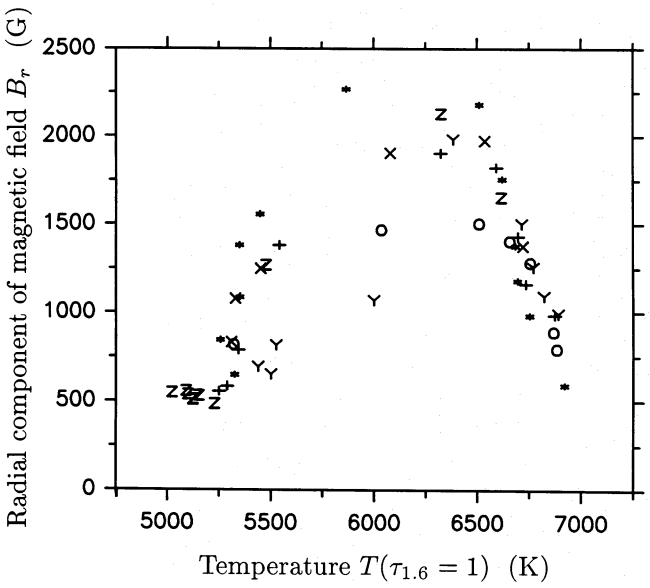

Fig. 5. Radial magnetic component, B_r, vs. $T(\tau_{1.6} = 1)$

Note that the rapid γ' change at the umbral boundary is far less obvious in the representation of γ' vs. position chosen in Paper V, although, with the help of hindsight, such a rapid change in γ' can be seen in the plots of the individual slices through the sunspot (Figs. 4d, 5d and 6d of Paper V). The rapid variation is smeared out and hidden by the variable umbral radius, r_u, in the standard representation of γ' vs. radial distance r from the centre of the spot (Fig. 8b of Paper V). This may explain why such anomalous behaviour of γ' has not been recognized earlier. Support for a rapid change of γ' at r_u is provided by the rapid change of the σ-to-π ratio of the 1.5648 μm line found by McPherson et al. (1992). Nevertheless, we cannot rule out a distortion of the γ' vs. T curve due to some hidden systematic error. Recall that γ' is less reliably determined than B. An independent confirmation of the γ'–T relationship, based on suitable data, would be welcome.

2.2. *Effects due to stray light*

Stray light can affect the relationships between B, γ', etc. and T. In Paper V we estimated the non-magnetic stray light in the penumbra and at some positions in the umbra directly from the fits to the line profiles (it was a free parameter of the inversion). The I_c that we use is already corrected for this straylight component. Except very near the outer penumbral edge, the non-magnetic stray light was found to contribute less than 5–10% to the total signal. The small amount of stray light is in keeping with the benign atmospheric conditions during the observations (granulation was clearly visible throughout) and the large and regular size of the spot (see Brinkle & Mattig 1965).

In addition to the non-magnetic stray light, there is also stray light coming from within the sunspot (magnetic stray light). Over most of the spot this component of the stray light simply leads to an averaging and a consequent loss of spatial resolution

(e.g., we average over most of the fine structure in the penumbra). The magnetic stray-light component is most important near the umbral boundary (e.g. Birkle & Mattig 1965) and we discuss its influence there. On the umbral side of the boundary, stray light from the penumbra enhances the continuum intensity. It may slightly shift the σ-components of the line (mimicking a slightly lower field strength) and it increases the π-to-σ ratio, making the field appear more horizontal. On the penumbral side of the boundary, the continuum intensity suffers from losses into the umbra, while the line profile shape is affected most by stray light from the nearby penumbra. In Figs. 2–5 the largest effect due to stray light (and finite spatial resolution) is expected for the points with 5500 K $\lesssim T(\tau_{1.6}=1) \lesssim$ 6500 K. Since for these points the error in T is much greater than in B, we expect stray light to shift points mainly in a horizontal direction in these figures. Under perfect conditions we would expect the number of points between 5500 K and 6500 K to be greatly reduced, but the plateau at approximately 2400 G in Fig. 2 is expected to remain and possibly even be extended towards higher and lower temperatures. The removal of stray light should decrease γ' near the umbral boundary, but to first order should not affect the large change in γ' visible in Fig. 3 between 5500 K and 6500 K. Thus the two main features of the observations at the umbral boundary are not expected to be greatly affected by stray light.

It remains to be shown that the spatial variations of the umbral brightness is real and not an artifact of magnetic stray light. In the umbra Martinez Pillet & Vazquez (1993), who take great pains to remove stray light, find a very similar relation between B and T. In particular, they find sizable variations in T within the umbra. The very high resolution images of Livingston (1992) also show considerable brightness structure within the umbra. We do not expect our sunspot to differ qualitatively from the others. Another question, namely why the darkest point (which does *not* lie in the centre of the umbra) is associated with the largest field strength, cannot be answered convincingly unless the umbra is inhomogeneous in brightness (the inhomogeneity in the magnetic splitting in the umbra is much too large to be explained away by stray light). Also note that even without a stray-light correction, the darkest point in the umbra is cooler than the coolest of the 3 Maltby et al. (1986) models describing the dark cores of sunspots.

In addition to the continuum intensity we also have 2 other diagnostics of the temperature in our data set, namely the CN 15646.2 Å line and the OH doublet at 15650.7 Å. Both lines show considerable variations across the sunspot and in particular within the umbra (see Figs. 4–6 of Paper V). While the depth of the OH line is beautifully anticorrelated with T in the umbra, the CN line shows a direct correlation. This contrasting behaviour again cannot be explained in terms of stray light, but only by a true variation of the temperature in the umbra.

Finally, there is the question whether the temperature variation in the penumbra is real or is an artifact of stray light. This is difficult to judge objectively, but its influence on the results of the present paper is not significant. The temperature enters into the analysis only through the gas pressure (Sect. 3) which is not strongly affected by the temperature variation in the penumbra (Fig. 8). Thus the uncertainty introduced into the penumbral results by the uncertainty in the amount of stray light is not expected to be larger than the quoted errors ($\pm$50 km uncertainty in $Z_{\rm W}$ due to measurement errors over most of the spot). We expect that the main uncompensated influence of the stray light on the results is to smear out the thermal boundary of the umbra and to increase the uncertainty of $Z_{\rm W}$ there. This must be borne in mind when considering the results of the present analysis.

3. Interpretation

3.1. Previous work

The first step of the derivation of a theoretical relationship between the magnetic vector and T is the integration of the radial component of the MHS force-balance equation in cylindrical symmetry (see Maltby 1977 for a clear derivation and a list of assumptions). This step gives a relationship between B_z, B_r and gas pressure P,

$$\begin{aligned} &P_0(z) - P(r,z) \\ &= \frac{1}{8\pi}\left(B_z^2(r,z) + 2\int_r^a B_z(r',z)\frac{\partial B_r(r',z)}{\partial z}\,dr'\right) \\ &= B_z^2(r,z)/8\pi + F_c(r,z)/8\pi. \end{aligned} \tag{3}$$

Here r and z denote the radial and vertical coordinate, respectively (with $r = 0$ corresponding to the symmetry axis of the sunspot), and a is the radial distance of a point that lies outside the sunspot (assumed to be in the quiet sun). In practice, one can take $a > r_p$, where r_p is the radius of the visible sunspot. F_c represents the curvature integral. Note that Eq. (3) is valid only at a constant height, z. Following the convincing arguments of Maltby (1977) we neglect velocity terms (measured Evershed velocities are always clearly subsonic).

The gas pressure within the sunspot, $P(r,z)$, may be eliminated by introducing the equation of state (ideal gas law), giving

$$\frac{T(r,z)}{T_0(z)} = \frac{m(r,z)}{m_0(z)}\frac{\rho_0(z)}{\rho(r,z)}\left[1 - \frac{B_z^2(r,z) + F_c(r,z)}{8\pi P_0(z)}\right], \tag{4}$$

where ρ is the gas density and m the mean molecular weight.

Early investigators simply pointed out the qualitative agreement in the umbra between theoretical prediction ($T \sim B_z^2$) and umbral observations ($T \sim B^2$). Martínez Pillet & Vázquez (1990) proposed that Eq. (4) can be used to obtain an estimate of the Wilson depression, $Z_{\rm W}$, which we define as

$$Z_{\rm W}(r,\lambda) = z(\tau_{0.5}=1, r > r_p) - z(\tau_\lambda = 1, r) = -z(\tau_\lambda = 1, r),$$

where τ_λ is continuum optical depth at wavelength λ for vertical incidence and z increases in the direction of decreasing τ. Following the usual practice we set $z(\tau_{0.5}=1) = 0$ in the quiet sun. Unless a wavelength is explicitely specified in the following, $Z_{\rm W}$ signifies the Wilson depression at 1.6 μm.

The basic idea is to make use of the fact that Eq. (4) is valid for a fixed z. If, using Eq. (4), the gas pressure in the quiet sun, $P_0(z)$, can be determined from the B and T measured in

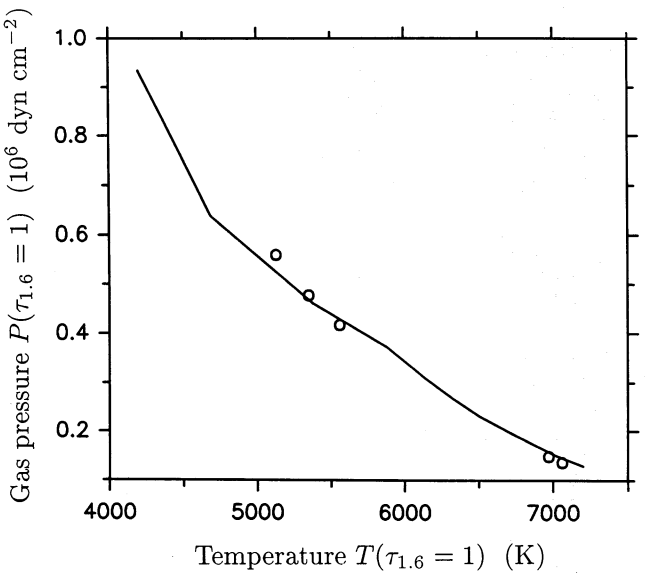

Fig. 6. Gas pressure at the continuum-forming layer, $P(\tau_{1.6}=1)$, vs. $T(\tau_{1.6}=1)$. See text (Sect. 3.2) for details

the sunspot, then a comparison of P_0 with a standard quiet-sun model can be used to fix the height at which B and T are determined. Martínez Pillet & Vázquez (1990, 1992) first rewrite Eq. (4) slightly:

$$\frac{T(r,z)}{T_0(z=0)} = \frac{m(r,z)}{m_0(z)}\frac{\rho_0(z)}{\rho(r,z)}\frac{T_0(z)}{T_0(z=0)} \times \left(1-\frac{1+f(r,z)}{8\pi P_0(z)}B^2(r,z)\right) , \quad (5)$$

where

$$f(r,z) = \frac{1}{B^2(r,z)}\left(F_c(r,z) - B_r^2(r,z)\right) . \quad (6)$$

Restricting themselves to the umbra, they simplify Eq (5) further by assuming that

$$\frac{m(r,z)}{m_0(z)}\frac{\rho_0(z)}{\rho(r,z)}\frac{T_0(z)}{T_0(z=0)} = \text{const} \quad (7)$$

and

$$f(r,z) = \text{const}$$

for all points in the umbra. Then, for a prescribed value of f, a regression is made through all (B,T)-pairs in the umbra, from which $P_0(z)$ may be deduced. By looking up the z at which the gas pressure in a tabulated quiet-sun convection-zone model (e.g. that of Spruit 1977) equals the derived P_0, an estimate of Z_W is finally obtained.

Martínez Pillet & Vázquez (1990, 1992) and Kopp & Rabin (1992) obtained reasonable values of Z_W with this technique for $0 \le f \le 1$. We have also applied it to our umbral data and find Z_W values between 470 and 1000 km for $-0.5 \le f \le 1$.

Martínez Pillet & Vázquez (1992) argue that variations in m, ρ, T_0, etc. do not invalidate their approach; only the derived Z_W becomes an average Wilson depression of the umbra. Their argument may well be valid in the umbra. Outside the umbra, however, we face a much larger range of Z_W values and the approach just outlined cannot be used without making additional, unsupported assumptions. This robs the 1.56 μm spectra of their main advantage, namely the ability to extend the B–T relationship to the whole sunspot.

It is relatively straightforward to resolve this problem if Z_W is determined individually for each spatial position in the sunspot. This approach, outlined in Sect. 3.2, has the additional advantage that it allows us to estimate the shape of the Z_W surface over the sunspot.

3.2. An improved procedure

To determine a stable value of Z_W throughout the sunspot we revert to Eq. (3). For known B_z, $P(r,z)$ and F_c a unique Z_W results from the fact that horizontal force balance is only satisfied for a given P_0, which is a unique function of z. The main uncertainties when applying Eq. (3) are caused by the necessity of choosing $P(r,z)$ and the curvature integral F_c. $P(r,Z_W)$ may be determined quite accurately from the measured $T(r,z)$ with the help of atmospheric models. The tension integral, F_c, is more of a problem, since it is non-local and requires a knowledge of B_z and B_r at a range of heights over much of the sunspot radius in order to be evaluated. For simplicity, we initially assume $F_c(r,z)=0$, determine $Z_W(r)$ from Eq. (3) and only then try to gauge the effects of a non-vanishing F_c on Z_W. We find that by comparing the $Z_W(r)$ determined here with the results derived from investigations of the Wilson effect we can set limits on F_c, $\frac{\partial B_r}{\partial z}$ and $\frac{\partial \gamma'}{\partial z}$ (Sect. 4.3). Equation (3) has the considerable advantage over Eqs. (4) and (5) that we need not prescribe $\rho_0(Z_W)/m_0(Z_W)$, which is strongly depth dependent.

We estimate $P(r,z)$ by using model atmospheres of different effective temperatures, $T_{\rm eff}$: We identify $P(r,z)$ with the $P(\tau_{1.6}=1)$ of an atmosphere whose $T(\tau_{1.6}=1)$ equals the $T(r,z)$ derived from the observations. The densest grid of model atmospheres which covers the full measured temperature range is that of Kurucz (1991). We use his set of atmospheres with logarithmic gravitational acceleration 4.5, solar abundances, 3500 K $\le T_{\rm eff} \le$ 6000 K, and a microturbulence of 2 km s^{-1}. Two neighbouring models in the grid are separated in effective temperature by 250 K. Although originally meant to describe stellar atmospheres, the Kurucz models are plane-parallel radiative and convective equilibrium models based on very detailed opacities that include the effects of millions of spectral lines and should adequately represent the atmospheric structure in sunspots.

To test whether the $P(\tau_{1.6}=1)$ values of the Kurucz models are acceptable for sunspots, we have compared them with the $P(\tau_{1.6}=1)$ values obtained from empirical solar quiet-sun and umbral models. The result of this comparison is plotted in Fig. 6. The solid curve connects the Kurucz values. The two circles near 7000 K represent the quiet sun models of Maltby et al. (1986) and Vernazza et al. (1976), the latter as modified in its lower layers by Spruit (1977). The circles between 5000 and 6000 K represent the umbral models E, M and L of Maltby et al. (1986). Judging from this figure, the Kurucz models are

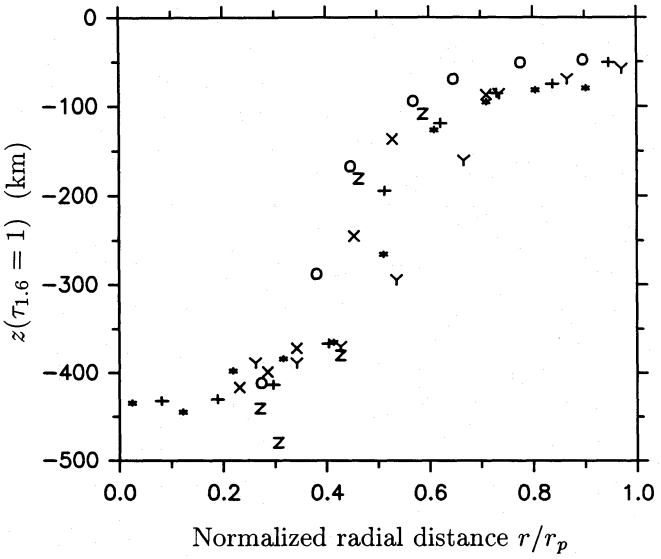

Fig. 7. $z(\tau_{1.6} = 1)$ vs. radial distance from the centre of the sunspot normalized to the outer penumbral radius, r/r_p. Symbols are the same as in Fig. 2

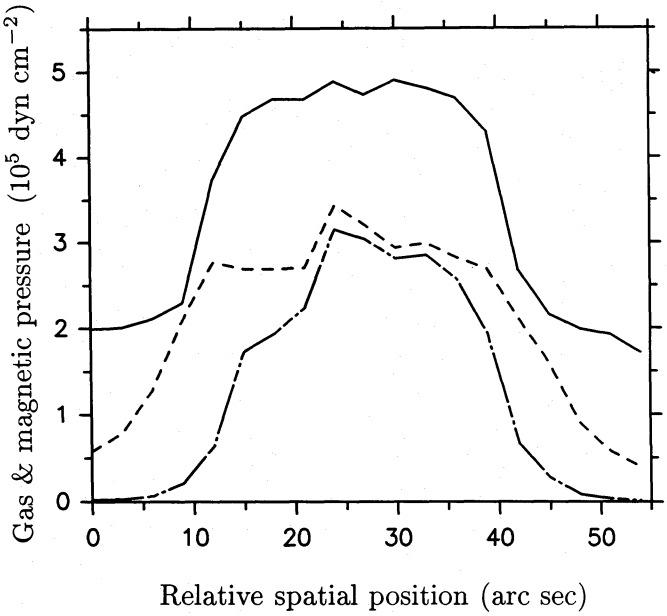

Fig. 8. Spatial variation of gas pressure, P (solid curve), magnetic pressure, $B^2/8\pi$ (dashed curve), and pressure due to the vertical component of the magnetic field, $B_z^2/8\pi$ (dot-dashed curve)

adequate for the present purposes. For any measured $T(\tau_{1.6} = 1)$ it is now possible to determine the corresponding $P(\tau_{1.6} = 1)$ by interpolating between the $P(\tau_{1.6} = 1)$ of the two Kurucz atmospheres closest in $T(\tau_{1.6} = 1)$ to the measured value.

Once $P(r, z)$ has been deduced, e.g. from Fig. 6, it is straightforward to determine $P_0(z)$ from Eq. (3) for each observed position within the sunspot and finally to obtain Z_W by looking up the z value in Spruit's (1977) convection-zone model corresponding to the deduced P_0.

4. Results

4.1. The Wilson depression

In Fig. 7 $z(\tau_{1.6} = 1) = -Z_W$ is plotted vs. r/r_p, the radial distance from the geometrical centre of the spot, normalized to the outer penumbral radius, r_p. The edge of the umbra, r_u, lies between $r/r_p = 0.4$ and 0.5. At least some of the scatter is solar. For example, the deepest point ('Z' symbol at $r/r_p = 0.3$) corresponds to the darkest part of the umbra, which also has the largest field strength. Furthermore, half-slices that have particularly high or low B values (e.g. 'Y' symbols or circles, see Fig. 8a of Paper V) exhibit particularly large or small Z_W values, respectively. $F_c \equiv 0$ was assumed when deriving the plotted Z_W values. Note that in the quiet sun $\Delta z = z(\tau_{0.5} = 1) - z(\tau_{1.6} = 1) \approx 20$ km, while in the umbra $\Delta z \approx 50$–60 km. Consequently, in the visible the Wilson depression is smaller than at 1.6 μm. We estimate that the total error in the derived Z_W due to measurement errors is of the order of 50 km for most points, but reaches 100 km near r_u, due mainly to stray light. The main uncertainty is due to the neglected curvature integral, F_c, which is discussed in Sect. 4.3.

4.2. Relative importance of gas and magnetic pressure

Let us briefly compare $P(\tau_{1.6} = 1)$ with the pressure due to the measured B and B_z, i.e. $B^2/8\pi$ and $B_z^2/8\pi$, respectively. All three pressures are plotted vs. spatial position in Fig. 8 for slices 1. The most striking feature of Fig. 8 is that the gas pressure is larger than the magnetic pressure at all positions in the continuum-forming layers of the sunspot. The plasma $\beta = 8\pi P/B^2$ drops from 1.4–1.8 in the umbra to nearly unity at the umbral boundary [which coincides with the sharp drop in $P(\tau_{1.6} = 1)$], then rises again to a value of 2–3 at the boundary of the sunspot. The average β of approximately 1.5–2 is in stark contrast to the average β of ≈ 0.3 in magnetic elements (Rüedi et al. 1992, Paper III of the present series).

Figure 8 also illustrates the relative importance of $P(\tau_{1.6} = 1)$ and B_z for determining Z_W. $P(\tau_{1.6} = 1)$ varies across the spot by nearly the same amount as $B_z^2/8\pi$ and $B^2/8\pi$, so that it is important to consider both the magnetic and the gas pressure when deriving Z_W.

The large $P(\tau_{1.6} = 1)$ gradient at the umbral boundary is due to the large temperature sensitivity of the H^- opacity. This change in gas pressure alone causes Z_W to change by approximately 200 km at r_u. Finally, Z_W does not go smoothly to zero at the outer penumbral edge (Fig. 7); there is another 'jump' of Z_W at r_p which is due to the decreased penumbral temperature, compared to the quiet sun.

4.3. Curvature forces and vertical gradients

The great unknown in the present analysis is the curvature integral, F_c. We have insufficient knowledge of $\frac{\partial B_r}{\partial z}$ to determine F_c independently and reliably judge its effect on Z_W. We can, however, turn the problem around and estimate F_c, $\frac{\partial B_r}{\partial z}$ and

$\frac{\partial\gamma'}{\partial z}$ by comparing the Z_W derived here with Z_W values obtained from Wilson-effect measurements made near the limb. It is valid to compare our results, obtained near the centre of the solar disc, with Wilson-effect observations near the limb, since the difference in z between $\tau_{0.5} = 1$ and $\tau_{0.5} = 0.1$ is 140–150 km in the quiet sun and approximately 100 km in the umbra.

4.3.1. F_c in the umbra

First consider the umbra. The Wilson effect gives $Z_W = 600 \pm 200$ km in the visible (Gokhale & Zwaan 1972; Zwaan, private communication 1992). Consequently, the average umbral Z_W derived by us, approxmately 400 km in the visible, lies at the lower end of the Z_W values allowed by Wilson-effect measurements. Thus, a negative $F_c(r = 0, Z_W(r = 0))$ can be ruled out, since it would decrease Z_W further. A tighter lower limit to $F_c(r = 0, Z_W)$ can be set by imposing a lower limit of 500 km on $Z_W(r = 0)$ in the visible following Maltby (1977), whose argument is based on a discussion of the pressure stratification of empirical models. The requirement that $Z_W \lesssim 800$ km in the visible (Gokhale & Zwaan 1972) gives an upper limit. Combining these two constraints we obtain

$$3.5 \times 10^5 \text{ dyn cm}^{-2} \lesssim F_c/8\pi \lesssim 1.6 \times 10^6 \text{dyn cm}^{-2}, \quad (8)$$

with a most probable value of 6×10^5 dyn cm^{-2}. Thus, in the umbra the curvature integral is of a similar magnitude as the gas and magnetic pressure terms (cf. Fig. 8), in agreement with the conclusions of Martínez Pillet & Vázquez (1992). A positive F_c implies that at the depth $Z_W(r = 0)$, the modulus of B_r averaged over the sunspot increases with height. To obtain a more quantitative estimate of this gradient, we assume that $\frac{\partial B_r}{\partial z}$ is independent of r at a given height. Then we can write

$$F_c = 2\frac{\partial B_r}{\partial z}\int_0^a B_z\, dr.$$

Since $\int_0^a B_z\, dr$ is the height independent magnetic flux of the sunspot, limits on $\left\langle \frac{\partial |B_r|}{\partial z} \right\rangle_{\text{spot}}$, i.e. $\frac{\partial |B_r|}{\partial z}$ averaged horizontally from $r = 0$ to $r = r_p$, can easily be set. An upper limit is found to be 0.3 G/km, with 0.07 G/km being a lower limit. If we assume $\frac{\partial B_r}{\partial z} \neq 0$ only in the penumbra, then $0.2 \lesssim \left\langle \frac{\partial |B_r|}{\partial z} \right\rangle_{\text{pen}} \lesssim 0.8$ G/km.

4.3.2. F_c and $\frac{\partial\gamma'}{\partial z}$ in the penumbra

Next we constrain $\frac{\partial\gamma'}{\partial z}$ in the penumbra. Following Maltby (1977) we rewrite $F_c(r, z)$ in terms of $\frac{\partial\gamma'}{\partial z}$,

$$F_c(r, z) = B_r^2 + 2\int_r^a \left(B^2 \frac{\partial\gamma'}{\partial z} - \frac{B_r^2}{r} \right) dr. \quad (9)$$

Note that in the absence of return flux (which, according to Paper V and Solanki & Schmidt 1992 cannot constitute a significant fraction of the total flux), $0 \leq \gamma' \leq \pi/2$. Equation (3) becomes

$$P_0(z) - P(r, z) = \frac{B^2(r, z)}{8\pi} + \frac{1}{4\pi}\int_r^a \left(B^2(r, z)\frac{\partial\gamma'(r, z)}{\partial z} - \frac{B_r^2(r, z)}{r} \right) dr. \quad (10)$$

Equation (9) suggests that the first and third terms of F_c tend to partially cancel each other. The B_r^2 term increases Z_W, while the $\int B_r^2/r\, dr$ term decreases Z_W. The influence of the $\int B^2 \frac{\partial\gamma'}{\partial z}\, dr$ term depends on the sign of the vertical γ' gradient. From the measured $B_r(r)$ dependence we can estimate the $-\int B_r^2/r\, dr$ term. At r_p it and the $\frac{\partial\gamma'}{\partial r}$ term vanish, so that, due to the B_r^2 term, Z_W is slightly larger at r_p than suggested by Fig. 7.[2] Taking the measured $B_r(r_p)$ value of 800 G we obtain that Z_W at r_p should increase by approximately 20 km relative to the values plotted in Fig. 7. In the inner penumbra ($r \approx 0.5 r_p$, $B_r \approx 2000$ G, $B \approx 2300$ G, $Z_W \approx 125$ km, $T(\tau = 1) \approx 6500$ K) we obtain

$$-\frac{1}{4\pi}\int_{r=0.5r_p}^a B_r^2/r\, dr \approx -1.5 \times 10^5 \text{ dyn cm}^{-2}. \quad (11)$$

This number almost completely compensates the first term of F_c in Eq. (9), $B_r^2/8\pi \approx 1.6 \times 10^5$ dyn cm^{-2}. Thus Z_W appears to be rather flat in the penumbra, if $\frac{\partial\gamma'}{\partial z}$ is neglected.

To estimate $\frac{\partial\gamma'}{\partial z}$ we assume that it is independent of r in the penumbra. Then, taking the measured radial dependence of B, we obtain

$$\frac{1}{4\pi}\int_{r=0.5r_p}^a B^2 \frac{\partial\gamma'}{\partial z}\, dr \approx 10^7 \left\langle \frac{\partial\gamma'}{\partial z} \right\rangle_{\text{pen}} \text{dyn cm}^{-2}, \quad (12)$$

where $\left\langle \frac{\partial\gamma'}{\partial z} \right\rangle_{\text{pen}}$ is expressed in units of °/100 km. Clearly, force balance requires that $\left| \frac{\partial\gamma'}{\partial z} \right|$ is well below unity (compare with Fig. 8). We set a lower limit on $\frac{\partial\gamma'}{\partial z}$ by requiring that the $\tau = 1$ level in the inner penumbra must not lie higher than in the outer penumbra, as dictated by Wilson-effect measurements. An upper limit is set by requiring that $Z_W \lesssim 650$ km (at 1.6 μm) at the inner penumbral boundary, since otherwise Z_W would become larger than 850 km in the umbra (recall that temperature alone causes Z_W to change by 200 km at r_u). These conditions give

$$-6 \times 10^{-3} \,°/\text{km} \lesssim \left\langle \frac{\partial\gamma'}{\partial z} \right\rangle_{\text{pen}} \lesssim 0.04 \,°/\text{km}. \quad (13)$$

Note that already for $\left\langle \frac{\partial\gamma'}{\partial z} \right\rangle_{\text{pen}} \lesssim -0.01°/\text{km}$, a static equilibrium cannot be maintained. Note too that when deriving the numbers in Eq. (13) we make use of the limits on $\frac{\partial B_r}{\partial z}$ (Sect. 4.3.1) and $\frac{\partial B}{\partial z}$ (Sect. 4.3.3). This allows us to constrain $\frac{\partial\gamma'}{\partial z}$ even if the $\tau = 1$ level in the penumbra is inclined.

4.3.3. F_c, Z_W and $\frac{\partial B}{\partial z}$ at the umbral boundary

Next we consider the influence of curvature forces on Z_W at r_u. According to Fig. 5, B_r changes by 500–1000 G at r_u over a radial distance of $\lesssim 4000$ km (Figs. 4, 5 and 6 of Paper V).

[2] Note that just inside r_p, the integral $\int B_z \frac{\partial B_r}{\partial z}\, dr$ does not vanish due to the large $\frac{\partial B_r}{\partial z}$ across the inclined boundary current sheet.

We can judge the influence of the curvature forces on Z_W by considering the integral

$$F_c(r_u - \Delta, z) - F_c(r_u + \Delta, z) = 2 \int_{r_u-\Delta}^{r_u+\Delta} B_z \frac{\partial B_r}{\partial z}\, dr, \quad (14)$$

which represents the curvature forces within a 2Δ thick layer around r_u. Choosing $\Delta = 2000$ km, $z = Z_W(r_u - \Delta)$ and $B_z = 2000$ G, we can set an upper limit on this integral by assuming that the whole observed variation of B_r across r_u is due to a vertical gradient $\frac{\partial B_r}{\partial z}$. In this case we see an apparent increase of B_r with r since we see down to different depths as we cross r_u. A lower limit follows from the assumption that the observed B_r variation is entirely due to a radial gradient of B_r, i.e. $\frac{\partial B_r}{\partial z} = 0$. The limits obtained in this manner appear reasonable. We find

$$0 \leq \frac{1}{4\pi} \int_{r_u-\Delta}^{r_u+\Delta} B_z \frac{\partial B_r}{\partial z}\, dr \lesssim 1.2 \times 10^6 \text{ dyn cm}^{-2}\,. \quad (15)$$

The lower limit corresponds to a 'jump' in Z_W of approximately 200 km and to $\frac{\partial B_r}{\partial z} = 0$. The upper limit implies that Z_W 'jumps' by approximately 500 km at r_u and that B_r increases with z by roughly 2.0 G/km. In order to obtain a unique value for the maximum possible $\Delta F_c = F_c(r_u - \Delta, z) - F_c(r_u + \Delta, z)$ and $\Delta Z_W = Z_W(r_u - \Delta) - Z_W(r_u + \Delta)$ we make use of the fact that F_c and Z_W are coupled together in two different ways. Firstly, they are coupled in an indirect manner through Eq. (3). The larger the $F_c(r_u - \Delta)$ (i.e. the larger the ΔF_c), the larger the $Z_W(r_u - \Delta)$ (i.e. ΔZ_W) becomes. Secondly, since B_r must change by around 1000 G over the height ΔZ_W, ΔZ_W determines $\frac{\partial B_r}{\partial z}$ at r_u and thus ΔF_c: The smaller the ΔZ_W, the larger the ΔF_c becomes. There is only a single pair of $(\Delta F_c, \Delta Z_W)$ values which satisfies both relations. Note that the maximum ΔZ_W and the additional constraint that $Z_W(r_u - \Delta) \lesssim 850$ km suggest that Z_W in the penumbra does not change too rapidly with r.

In principle, it is possible to eliminate the jump in Z_W at r_u by postulating $\frac{\partial B_r}{\partial r} \approx -1$ G/km in a 2000 km wide layer around r_u. However, in order to reproduce the classical picture of a sloping penumbral surface with a smooth transition of Z_W into the umbra, three distinct conditions must be simultaneously met:

1. $\frac{\partial B_r}{\partial z}$ must be positive in the penumbra in order to obtain a sloping $\tau = 1$ surface and a sufficiently large Z_W in the umbra.
2. Near r_u, $\frac{\partial B_r}{\partial z}$ must be negative and of just the right magnitude to compensate for the Z_W jump due to the temperature change at r_u.
3. Near r_u, $\frac{\partial B_r}{\partial r}$ must be positive and very large, compared to the rest of the sunspot [in order to explain the observed $B_r(r)$].

To postulate the exact behaviour of B_r implied by these three conditions appears artificial and unnecessarily complex. In the absence of independent evidence favouring the classical picture (Wilson effect measurements support a relatively sharp change in Z_W at r_u, e.g. Wilson & Cannon 1968), it must therefore be rejected. Consequently, we conclude that our analysis supports a jump of 200–500 km in Z_W at r_u.

Finally, the fact that a 'jump' is present at Z_W, but no significant change is seen in B at r_u implies that, as a conservative estimate, $|\frac{\partial B}{\partial z}| \lesssim 2$ G/km at r_u for a jump in Z_W of 200 km. If the jump is 500 km, then $|\frac{\partial B}{\partial z}| \lesssim 0.8$ G/km. These limits are in good agreement with the estimate of 0.4 G/km at r_u made by Beckers & Schröter (1969).

5. Conclusions

We have investigated the relationship between the temperature, T, and the magnetic vector (B, γ') using spectra at 1.56 μm. The $B(T)$ relationship found by Kopp & Rabin (1992) is confirmed. In addition, a linear relation between γ' and T is found.

By applying the radial component of the force balance equation to the observed relationship, we have obtained an estimate of the Wilson depression, Z_W, as a function of position in the sunspot. Z_W at r_p is approximately 40–60 km and appears to increase only slowly towards the umbra. At the umbral boundary, r_u, Z_W increases by another 200–500 km. This qualitative radial dependence of Z_W agrees well with the picture obtained by utilizing the Wilson effect (Wilson & Cannon 1968; Wilson & McIntosh 1969; Wittmann & Schröter 1969). Quantitative differences to the results of Wilson-effect measurements are due to our insufficient knowledge of magnetic curvature forces. The comparison between the $Z_W(r)$ derived here (≈ 400 km in the umbra) and that obtained from the Wilson effect (600 ± 200 km, Gokhale & Zwaan 1972) provides constraints on the curvature integral, $\frac{\partial B_r}{\partial z}$ and $\frac{\partial \gamma'}{\partial z}$. We find that the curvature force per unit area can be of the same magnitude as the gas and magnetic pressure.

Of particular interest are the tight limits we can set on $\frac{\partial \gamma'}{\partial z}$. Their most visible consequence concerns the explanation of the broad-band circular polarization proposed by Sánchez Almeida & Lites (1992). Their, in all other respects very promising, approach requires a rapid change of γ with optical depth, namely $\frac{\partial \gamma}{\partial \tau} = 45° \sin\gamma$ for $\gamma \leq 90°$, where γ is the angle between the magnetic vector and the line-of-sight. For a vertical line-of-sight $\gamma = \gamma'$, so that in the penumbra (i.e. $\sin\gamma' \gtrsim 0.8$) $36° \lesssim \left\langle \frac{\partial \gamma'}{\partial \tau} \right\rangle_{\text{pen}} \lesssim 45°$. Since a decade in τ corresponds to approximately 100–150 km in z, this model implies $\frac{\partial \gamma'}{\partial z} \approx -12$ to $-20°/100$ km in the lower photosphere of penumbrae, which is the relevant height range for the production of the broad-band circular polarization. The required gradient is over an order of magnitude larger than the limit of $-0.5°/100$ km set in Sect. 4.3. Thus, although a mechanism based on a change in γ with τ may well be responsible for the broad-band circular polarization observed by Illing et al. (1974a,b, 1975) and Makita & Ohki (1986), a different model from the one proposed by Sánchez Almeida & Lites (1992) is required to explain it in a physically consistent manner. An attempt to reconcile the limits

on a global $\frac{\partial \gamma'}{\partial z}$ with the observed broad-band circular polarization is made by Solanki & Montavon (1992).

The present analysis can be improved in three main aspects. 1. Most urgent are tighter independent constraints on the Wilson depression, e.g. from Wilson-effect measurements with high spatial resolution. The better such constraints, the narrower the limits become that can be set on curvature terms and magnetic gradients. 2. The validity and universality of the γ' vs. T relation should be tested by observing many more sunspots, if possible with Stokes vector data. 3. Once curvature terms have been better estimated, a 2-D image of B and T in a sunspot (derived from the g = 3, 1.56 μm line and the nearby continuum) should provide a 3-D picture of the Wilson depression.

Acknowledgements. We are greatly indepted to J.O. Stenflo for detailed discussions which helped to clarify many of our ideas. We thank P. Maltby for stressing the importance of stray light.

References

Alfvén H., 1943, Arkiv. Mat. Astron. Fysik 29, No. 11
Beckers J.M., Schröter E.H., 1969, Sol. Phys. 10, 384
Birkle K., Mattig W., 1965, Z. Astrophys. 60, 264
Cowling T.G., 1957, Magnetohydrodynamics, Interscience, New York (p. 21)
Dicke R.H., 1970, ApJ 159, 25
Gokhale, M.H. and Zwaan, C., 1972, Sol. Phys. 26, 52
Illing, R.M.E., Landman, D.A., Mickey, D.L., 1974a, A&A 35, 327
Illing, R.M.E., Landman, D.A., Mickey, D.L., 1974b, A&A 37, 97
Illing, R.M.E., Landman, D.A., Mickey, D.L., 1975, A&A 41, 183
Kopp G., Rabin D., 1992, Sol. Phys. 141, 253
Kurucz R.L., 1991, in Stellar Atmospheres: Beyond Classical Models, L. Crivellari, I. Hubeny, D.G. Hummer (Eds.), Kluwer, Dordrecht, p. 441
Lites, B.W., Skumanich, A., 1990, ApJ 348, 747
Livingston W., 1991, BAAS 23, 1030
Makita M., Ohki Y., 1986, Ann. Tokyo Astron. Obs. 21, 1
Maltby P., 1977, Sol. Phys. 55, 335
Maltby, P., Avrett, E.H., Carlsson, M., Kjeldseth-Moe, O., Kurucz, R.L., Loeser, R., 1986, ApJ 306, 284
Martínez Pillet V., Vázquez M., 1990, Ap&SS 170, 75
Martínez Pillet V., Vázquez M., 1993, A&A 270, 494
McPherson M.R., Lin H., Kuhn J.R., 1992, Sol. Phys. 139, 255
Rüedi I., Solanki S.K., Livingston W., Stenflo J.O., 1992, A&A 263, 323 (Paper III)
Sánchez Almeida J., Lites B.W., 1992, ApJ 398, 359
Solanki S.K., Montavon C.A.P., 1993, A&A 275, 283
Solanki S.K., Schmidt H.U., 1993, A&A 267, 287
Solanki S.K., Rüedi I., Livingston W., 1992a, A&A 263, 312 (Paper II)
Solanki S.K., Rüedi I., Livingston W., 1992b, A&A 263, 339 (Paper V)
Spruit, H.C., 1977, Ph.D. Thesis, Univ. Utrecht
Vernazza, J.E., Avrett, E.H., Loeser, R., 1976, ApJS 30, 1
Walther U., 1992, Diplomarbeit, ETH, Zürich
Wilson P.R., Cannon C.J., 1968, Sol. Phys. 4, 3
Wilson P.R., McIntosh P.S. 1969, Sol. Phys. 10, 370
Wittmann A., Schröter E.H., 1969, Sol. Phys. 10, 357