THE STRENGTH OF THE SUN’S POLAR FIELDS

LEIF SVALGAARD, THOMAS L. DUVALL, JR., and PHILIP H. SCHERRER

Institute for Plasma Research, Stanford University, Stanford, Calif. 94305, U.S.A.

(Received 6 April, 1978)

Abstract. The magnetic field strength within the polar caps of the Sun is an important parameter for both the solar activity cycle and for our understanding of the interplanetary magnetic field. Measurements of the line-of-sight component of the magnetic field generally yield 0.1 to 0.2 mT near times of sunspot minimum. In this paper we report measurements of the polar fields made at the Stanford Solar Observatory using the Fe I line λ525.02 nm. We find that the average flux density poleward of 55° latitude is about 0.6 mT peaking to more than 1 mT at the pole and decreasing to 0.2 mT at the polar cap boundary. The total open flux through either polar cap thus becomes about $3 \times 10^{14}$ Wb. We also show that observed magnetic field strengths vary as the line-of-sight component of nearly radial fields.

1. Introduction

The magnetic field strength within the polar caps of the Sun is an important parameter for both the solar activity cycle and for our understanding of the interplanetary magnetic field. Measurements of the line-of-sight component of the magnetic field generally yield 0.1 to 0.2 mT (e.g., Severny, 1971; Howard, 1974a) near times of sunspot minimum. Near sunspot maximum the field weakens and reverses polarity (Babcock, 1959; Howard, 1974a). At the present solar minimum, magnetic field strengths at latitudes above 60° are observed to be in the 0.0 to 0.2 mT range. In a recent study, Howard (1977) concludes that the ‘true’ field strength is unlikely to be more than twice the measured values, and that it is impossible that it could be greater than the measured line-of-sight values by more than a factor of five.

The true magnetic field strength is not the same as the average field strength measured by the solar magnetograph because the bulk of the magnetic flux is confined to sub-arc second bundles of field lines where field strengths are of the order of 100–200 mT (Stenflo, 1973). It would be more correct to acknowledge this fact by referring to an average flux density over some area on the Sun rather than to a field strength.

In this paper we report measurements of the polar fields made at the Stanford Solar Observatory using the Fe I line λ525.02 nm (Scherrer et al., 1977; Duvall, 1977). We find that the average flux density poleward of 55° latitude is about 0.6 mT peaking to more than 1 mT at the pole and decreasing to 0.2 mT at the polar cap boundary. The total open flux through either polar cap thus becomes about $3 \times 10^{14}$ Wb.

2. Observational Details

In the magnetograph mode the solar image is scanned with a $175'' \times 175''$ square aperture. The image is divided into 11 scan lines (Figure 1) parallel to the solar
equator. The aperture is stepped one-half aperture between each measurement. A typical error for both the magnetic signal and the zero level is 5 \mu T. Figure 2 shows a typical magnetogram. Three active regions all belonging to the new sunspot cycle – number 21 – are clearly seen. In addition, extensive background fields are apparent including well-defined polar fields. Referring to Figure 1 we shall be particularly interested in the field measured in the polar cap, i.e. the polemost scan line in each hemisphere. The average latitude of the equatorward limit of that aperture is 55°. At the present time that is also the equatorward boundary of the polar coronal holes (Solar Geophysical Data, 1976, 1977). In accordance with that, the measured polar fields have invariably been unipolar (over the aperture) throughout the interval covered by the daily magnetograms – 16 May, 1976 through August, 1977 – positive or outwards in the north and negative or inwards in the southern polar cap. The position of an aperture along the scan line is measured by the parameter x, being 0 at central meridian, and \pm 11 respectively at the west – and east equatorial limbs (Figure 1). We shall be concerned with apertures at x = 0 (black in Figure 1) and at x = \pm 2 (dotted areas in Figure 1). During the course of a year the solar rotation axis tips toward and later away from the observer by 71°. We thus get a kind of stereoscopic view of the polar fields provided that they do not change appreciably over a timescale of up to a year.
Fig. 2. Typical Stanford magnetogram (5 March, 1977). Solid contours denote positive fields while dashed contours denote negative fields. Three active regions are readily seen as well as extended areas of background fields. Contour levels are $\pm 0.02$, $\pm 0.05$, $\pm 0.1$, $\pm 0.2$, $\pm 0.5$, $\pm 1$, $\pm 2$ mT.

In order to interpret measurements of the polar cap fields, it is necessary to investigate if the fact that the data is taken near the limb is distorting the data in any way. It has been suggested by Howard and Stenflo (1972) that measurements of the magnetic field in the $\lambda 525.02$ nm line are systematically in error by a factor that depends on the angle between the radius and the line of sight. This error changes by more than a factor of two as one goes from the center of the disk to the limb. By following specific areas as they cross the disk from east limb to west limb in the course of solar rotation we may directly verify the existence of this systematic error under the weak assumption that the intrinsic field does not change considerably during disk passage. In the following section we report the results of an extensive analysis of measured field strengths as a function of the distance from disk center. Our conclusion is that no systematic error of the kind suggested by Howard and Stenflo (1972) is apparent in the Stanford data.

3. Center to Limb Variation of Field Strength

When an aperture is placed at central meridian an area is defined on the solar surface. The area can be assigned heliographic coordinates and can be identified on magnetograms both before and after central meridian passage. For the analysis we utilize the three near-equatorial scan lines – ranging up to $16^\circ$ in latitude on either side of the equator. If the dominant effect that determines the line-of-sight component, $B_t$, of the field $B$ is simple projection, the observed longitudinal field should vary as $B_t = B \cos L$ where $L$ is the angle between the radius vector and the line of sight (Figure 3). Figure 4 shows the average $B_t$ as a function of $\cos L$. The
Fig. 3. Definition of the angle $L$ and line-of-sight component $B_l$ for a radial field in the equatorial plane.

analysis was performed separately for weak fields ($|B| < 150 \mu T$) and for strong fields ($|B| \geq 150 \mu T$). The magnetic field strength at central meridian passage was used as the selection parameter. Furthermore, the data was analyzed separately for each polarity of the field. More than 90% of the data fall in the weak field category—which ultimately also includes the polar regions. It is difficult to assess the error bars in Figure 4 because the data is not independent but conservative estimates based on the standard deviation of the values going into each average point lead to standard errors of the mean of $\pm 1 \mu T$ for the weak field case and of $\pm 50 \mu T$ for the strong field case. It appears that the observed field strength is very nearly proportional to $\cos L$. There is a small asymmetry between east and west hemispheres: at the east limb a small net negative field is observed, while at the west limb a net positive field is seen. The imbalance is caused by a weak net azimuthal field first observed by Howard (1974b) and discussed in detail by Duvall (1977). Folding the west hemisphere onto the east hemisphere data eliminates this azimuthal field component as shown by the lower line in Figure 5. Also shown in Figure 5 is the result of a similar analysis of the latitude band $16^\circ - 32^\circ$. In this latitude band most of the active regions have appeared over the time interval covered by the analysis. The aperture at disk center corresponds to a certain area on the solar surface ($1.7 \times 10^{10} \text{ km}^2$); at some distance from disk center the area is larger. This increases the chances that more active region flux cancels inside the aperture and thus lowers the observed average field strength. This effect can be discerned in Figure 5 for the $16^\circ - 32^\circ$ latitude bands for small values of $\cos L$—corresponding to a larger area within the aperture.
Fig. 4. *Upper panel:* Observed center-to-limb variation of measured magnetic field $B_t$ for weak average fields (less than $150 \, \mu T$ at central meridian). Data within $16^\circ$ of the equator is used. Open symbols signify positive fields and filled symbols show the variation of negative fields. Each point is an average of more than 500 measurements. *Lower panel:* Same as above, but for strong fields only (greater than or equal to $150 \, \mu T$ at central meridian passage). The number of cases is much lower than for the weak field regions and the scatter is correspondingly larger. In addition, the assumption of constant intrinsic field strength for the two-week disk passage is more likely to be invalid for these strong fields.

The overall result seems to be that the measured magnetic field values fall off very nearly as $\cos L$ from central meridian to the limb, i.e., consistent with the interpretation of the observed field being just the line-of-sight component of the real field. By assuming that the field measured using Fe I $\lambda 523.3$ nm represents the true average field — unaffected by Zeeman saturation and line weakening — Howard and Stenflo (1972) suggested that the field as observed in $\lambda 525.0$ nm should be corrected by a factor $(0.48 + 1.33 \cos \rho)^{-1}$ where $\rho$ is the angle between line-of-
Fig. 5. Folding east-hemisphere data onto west-hemisphere data and negative field magnitudes onto positive field magnitudes results in a single curve showing field magnitude as a function of cos L from central meridian (L = 0°) to the limb (L = 90°). The filled circles show the result for the weak field case of Figure 4. Open circles show the result of a similar analysis for the latitude bands (16°–32°) in both hemispheres. Here all fields were included and the effect of flux cancellation within apertures closer to the limb can be discerned. The straight lines were forced to go through zero at the limb. The main result of Figures 4 and 5 is that we are simply measuring longitudinal magnetic fields.

sight and the radius vector. In Figure 6 we show the center-to-limb variation of a radial field of 51 μT as it would be observed at Stanford and as it should be observed if the correction factor suggested by Howard and Stenflo (1972) is applied. The linear dependence on cos L shown by the Stanford data (Figure 4) is clearly inconsistent with any such correction. The reason for this discrepancy is not known, but collaborative work is in progress to clear up the matter. One might speculate (Stenflo, 1977, personal communication) that since the field observed in the λ 523.3 nm line does not refer to the same height in the photosphere at the center of the disk and at the limb, a rapid divergence of the field lines in the upper photosphere where the λ 523.3 nm line is formed may distort the center-to-limb variation of the measured fields. In the deeper layers where λ 525.0 nm is formed
Fig. 6. Filled circles show how a magnetic field of 51 μT at central meridian would be observed at Stanford ($B_s$) to change towards the limb as a function of cos $L$. Assuming that $B_s$ is the true average longitudinal field, the field ($B_w$) that would be observed at Mount Wilson as obtained by applying the Howard and Stenflo (1972) correction is shown with open circles. Normalizing the $B_w$ curve to give the same field at disk center as observed at Stanford yields the dotted curve (crosses). It is clear that the Stanford data (Figure 5) is inconsistent with any systematic error of the kind proposed by Howard and Stenflo.

The divergence is probably insignificant. The verification of this requires detailed calculations with fluxtube models and multi-dimensional radiative transfer and is outside the scope of the present paper. At the present time, we adopt the simplest interpretation, namely that the low-noise Stanford magnetograms indicate that magnetic field strength values obtained in the $\lambda$ 525.0 nm line can be corrected for foreshortening by dividing with cosine of the angle between the radius vector and the line-of-sight.

These considerations have implications for the computation of potential magnetic fields in the corona. In the past Mt. Wilson $\lambda$ 525.0 nm data has been used with the correction suggested by Howard and Stenflo (1972), e.g. Riezebieter (1977), and Kitt Peak $\lambda$ 523.3 nm data have been used assuming no distortion of the limb data due to field line divergence, e.g. Altschuler et al. (1977). Quantitative results from these calculations, e.g. Levine et al. (1977), should now be considered suspect.
4. Magnetograph Saturation

At Stanford we use the same slit sizes and arrangement as at Mt. Wilson for the line $\lambda$ 525.02 nm. The two exit slits are $\lambda$ 7.5 pm wide and have a separation of $\lambda$ 1.8 pm. With these slits the magnetic signal as a function of effective Zeeman splitting is shown in Figure 7. The curve in Figure 7 is obtained by placing a right-hand circular polarizer in front of the KDP crystal and recording the magnetic signal while scanning across the line. For weak fields (i.e., less than 50 mT) the magnetic signal is proportional to the field strength. As the field increases the magnetograph response weakens and at 143 mT the magnetograph is saturated and any further increase actually decreases the magnetic signal. Assuming that a typical field strength within the magnetic elements is 150 mT (Stenflo, 1973; Ramsey et al., 1977) the corresponding reading from the magnetograph would be only 83 mT; hence the effect of magnetograph saturation due to the strong fields in the elements is to reduce the measured flux by a factor 150/83 = 1.8.

![Diagram showing calibration curve for the Stanford magnetograph for the Fe I line $\lambda$ 525.02 nm and for exit slits $\lambda$ 7.5 pm wide separated by $\lambda$ 1.8 pm. A magnetic field produces a Zeeman splitting of $\Delta\lambda$ = 38.6 pm/T. This relation is shown as the dotted line and is also used to calibrate the scale along the upper right boundary of the figure frame. Slit dimensions are shown in the middle of the figure.](image)

The careful, lengthy analysis by Stenflo (1973) gives essentially the same result. If the polar fields also exist mainly in the form of elements of the same strength as in lower latitudes we might expect a similar reduction of the measured flux. No detailed study of this problem has been carried yet specifically within the polar caps but preliminary data from the Stanford Solar Observatory suggests that there is no difference in the saturation between the polar and the equatorial limbs.
5. Polar Field Strengths

In order to set the stage for the discussion of the Stanford data, Figure 8 shows the average line-of-sight magnetic field strength poleward of $69^\circ$ latitude (filled symbols) and between $60^\circ$ and $69^\circ$ (open symbols) observed at Mt. Wilson since 1971 following the polar field reversal around 1970. Average field strengths polewards of $55^\circ$ as observed at Stanford during 1976 and 1977 appear as symbols with crosses and seem to concur with the Mt. Wilson averages. Assuming that the polar fields are radial and that an effective latitude of the polar cap is $70^\circ$ we may arrive at a first crude estimate of the true field strength

$$B_{true} = (\pm 100 \mu T / \cos 70^\circ) \times 1.8 = 0.5 \text{ mT}.$$  

The Stanford data allows us to extend the analysis considerably. If the polar fields were indeed uniform and radial they would not change during the year as the rotation axis tips back and forth by $71^\circ$. Figure 9 shows that the polar fields change by about a factor of two during the year. We have plotted monthly averages of the field strengths through polar scan line apertures $x = 0$ and $x = \pm 2$ for both hemispheres. There can be no doubt about the reality of the apparent yearly variation shown in Figure 9. The uncertainty in each average point is less than 10 $\mu$T and is
Fig. 9. Yearly variation of magnetic field strength averaged over poleward apertures $x = 0$ (upper panel) and $x = \pm 2$ (lower panel) observed at Stanford May 1976–August 1977. Times of maximum heliographic latitude of the Earth are marked by vertical lines and arrows. Open circles represent data from May–August 1976 that are plotted as May–August one year later. Each data point is a monthly average.

mainly due to solar causes rather than instrumental errors. Times of maximum tilt away from the observer are indicated by arrows while times of maximum tilt towards the observer are marked by short vertical lines.

To proceed we now introduce a model of the polar field geometry. Judging from the inclination of polar plumes in eclipse photographs, it seems likely that the magnetic field lines at high latitudes are inclined toward the equator. We therefore introduce a meridional field component, $B_\theta$, arising from a field inclination $\alpha$ against the radial, (Figure 10). The radial component is modeled by the expression

$$B_r = B_p \cos^n \theta$$

and also

$$B_\theta = B_r \tan \alpha ,$$

where

$$\alpha = q \theta .$$
Here \( n \) and \( q \) are the two parameters of the model with \( B_p \) providing the calibration being the field strength at the very pole. \( n = 0 \) is the case of a uniform radial component of the field across the polar cap, while a large value of \( n \) indicates a field distribution highly peaked towards the pole. Similarly \( q = 0 \) corresponds to a purely radial field, while \( q = 1 \) is typical of the inclination of the polar plumes. A dipole field has \( q = 0.5 \) within the polar cap. By varying \( n \) and \( q \) a wide range of field geometries can be accommodated.

The next step is to calculate the average line-of-sight component of the model field properly weighted by limb darkening within each of our apertures for various values of polar axis tilt angle. Of special interest are the two cases of maximum tilt away from and towards the observer. In the model calculations \( B_p \) is set equal to one. Let the computed average be \( \langle B \rangle_{\text{comp}} \) and the observed value be \( B_{\text{obs}} \), then \( B_p = B_{\text{obs}}/\langle B \rangle_{\text{comp}} \) is our calibration constant. The ratio between the average model fields through a given aperture for the pole tipped toward us and for the pole tipped away from us by the maximum amount, i.e.

\[
r = \frac{\langle B \rangle_{\text{comp}}[\text{toward}]}{\langle B \rangle_{\text{comp}}[\text{away}]}
\]

is a simple measure of the non-uniformity of the magnetic fields within the polar cap. This ratio was observed to be about \( 2^* \). Figure 11 shows the computed ratio as

* A more detailed analysis gives \( r = 2.1 \pm 0.1 \) for both apertures shown in Figure 9.
a function of $n$ and $q$. The solid contours are for the aperture $x = 0$, while dashed contours are for the case $x = \pm 2$. It is fortunate that the contours for the two cases do not run nearly parallel to each other. As things stand it seems possible to conclude that $n = 8 \pm 1$ and $q = 0.0 \pm 0.1$ fits the observed value of $r = 2.1$ uniquely. The resulting $B_p$ turns out to be 0.64 mT. The variation of $B$ with latitude is shown in Figure 12. The ordinate scale on the left is for field strengths observable with $\lambda 525.02$. On the right the effect of magnetograph saturation is incorporated by scaling with the factor 1.8. We see that the data is consistent with rather high field strengths at the pole ($B_p = 1.15$ mT) falling rapidly off with colatitude until about 0.2 mT at the polar cap boundary. It is essentially this strong peaking of the field strength near the pole that causes the large change of line-of-sight field with time of year as seen in Figure 9.

The polar cap magnetic flux is given by

$$F = R^2_\odot \int_0^{\theta_c} \int_0^{2\pi} B(\theta) \sin \theta \, d\lambda \, d\theta = 2\pi R^2_\odot B_p \frac{1}{3}(1 - \cos^9 \theta_c),$$

where $\theta_c$ is the colatitude of the polar cap boundary. Because the area of the polar cap is $A = 2\pi R^2_\odot (1 - \cos \theta_c)$ the average field strength becomes

$$\langle B \rangle_{\text{cap}} = B_p \frac{1}{9} \frac{1 - \cos^9 \theta_c}{1 - \cos \theta_c} = 0.51B_p$$

$$= 0.65 \text{ mT}.$$
It is worth pointing out that the polar cap flux, \( F = 3.24 \times 10^{14} \text{ Wb} \), is almost sufficient to provide the interplanetary magnetic flux (\( F_{\text{IMF}} = 5 \times 10^{14} \text{ Wb} \) for each hemisphere). With additional flux entering the solar wind from low latitude magnetic sectors it seems that the interplanetary magnetic flux is easily accounted for at least when the solar fields are strong near sunspot minimum. At times shortly after sunspot maximum the equatorial sectors must supply all the necessary flux to maintain the observed constancy of the interplanetary magnetic field (e.g. King, 1976).

6. Coronal Holes

It is often assumed (e.g., Hundhausen, 1977; Levine et al., 1977; Suess et al., 1977) that coronal holes have unique properties independent of their position on the Sun; in particular that the photospheric magnetic field strength is always the same within coronal holes, namely about 1 mT. During the interval of our observations (May 1976 – August 1977) there were only a few and ill-defined equatorial holes (Solar Geophysical Data, 1976, 1977). The polar caps had however prominent extensions.
or lobes towards lower latitudes. These extensions are long lived and occur near the middle of magnetic sectors (e.g., Wilcox and Svalgaard, 1974). A particularly good example can be found in the northern hemisphere around heliographic longitude 275° during Carrington rotation 1651–1653. The magnetic field within the hole was rather constant and is shown in Figure 12 as open circles. It appears that this coronal hole magnetically can be described as an extension of the polar cap; the field strength of 0.15 mT being typical of polar cap boundary conditions. It would seem that physical models of coronal holes would benefit from the more realistic estimates of the magnetic field strength suggested in the present paper.

In a numerical simulation of conditions within the northern polar hole during July 1973 from \(2R_\odot\) to \(5R_\odot\), Suess et al. (1977) show that the dominant source of divergence within \(5R_\odot\) is the meridional gradient of the magnetic field strength. The resulting meridional magnetic pressure gradient relieves itself through meridional plasma flow producing the observed large spreading of the polar hole with altitude. The magnetic field geometry deduced in the present paper is entirely consistent with these ideas.

7. Summary

Observations at Stanford Solar Observatory of solar magnetic fields in the Fe I line \(\lambda 525.02\) nm shows that a radial magnetic field measured at a point where the radius makes an angle \(\rho\) with the line of sight is observed to be decreased by a factor \(\cos \rho\). For field elements of 150 mT, magnetograph saturation causes the measured field to be too low by a factor of 1.8. The average field magnitude poleward of 55° latitude is measured to be near 100 \(\mu\)T. Variation of the apparent field over a 3 arc min aperture grazing the limb at central meridian amounts to a factor of two over the year; the field being strongest when the pole is tipped the most (74°) towards the observer. Combination of all the above results leads to the following picture of the magnetic field within the polar caps. The field is nearly radial, varying as \(B_\rho \cos^8 \theta\) where the field strength \(B_\rho\) at the pole \((\theta = 0°)\) is 1.15 mT, and falling off to below 0.2 mT at the polar cap boundary \((\theta = 35°)\). Within coronal holes outside of the polar cap the magnetic field strength at sunspot minimum is rather small (0.15 mT). At other phases of the sunspot cycle the relative strengths of polar and lower latitude field magnitudes is very likely different from what is observed at the present sunspot minimum.

Acknowledgements

We thank J. M. Wilcox, R. Howard, and J. O. Stenflo for extensive discussions and R. Howard for the use of Mt. Wilson data. Most of the observations at Stanford were performed by E. Gustafson and S. Bryan. We have benefited greatly from the stimulus provided by the Skylab workshop on coronal holes. The present work
was supported in part by the Office of Naval Research under Contract N00014-76-C-0207, by the National Aeronautics and Space Administration under Grant NGR 05-020-559, by the Atmospheric Sciences Section of the National Science Foundation under Grants ATM74-19007 and DES75-15664, and by the Max C. Fleischmann Foundation.

References