SOLAR MAGNETIC CYCLE

KAREN L. HARVEY
Solar Physics Research Corporation, 4780 Calle Desecada Tucson, AZ 85718,
U.S.A.

Abstract. Using NSO/KP magnetograms, the pattern and rate of the emergence of magnetic flux and the development of the large-scale patterns of unipolar fields are considered in terms of the solar magnetic cycle. Magnetic flux emerges in active regions at an average rate of \(2 \times 10^{21} \text{ Mx/day} \), approximately 10 times the estimated rate in ephemeral regions. Observations are presented that demonstrate that the large-scale unipolar fields originate in active regions and activity nests. For cycle 21, the net contribution of ephemeral regions to the axial dipole moment of the Sun is positive, and is of opposite sign to that of active regions. Its amplitude is smaller by a factor of 6, assuming an average lifetime of ephemeral regions of 8 hours. Active regions larger than 4500 Mm\(^2\) are the primary contributor to the cycle variation of Sun's axial dipole moment.

Key words: solar cycle, magnetism, active regions

1. Introduction

The solar cycle is produced and modulated by the emergence and interaction of magnetic fields, leading to an array of activity over a large range of spatial and time scales, all related and the result of as yet unknown processes in the solar interior. Our knowledge of the solar cycle comes from observations and the interpretation of the distribution and evolution of magnetic fields at or above the photosphere. Activity belonging to any given cycle can be observed for approximately 15 years, while successive cycles begin at intervals of around 11 years. Over the course of a solar cycle, the patterns of magnetic fields in the photosphere vary in a systematic way.

During the three to four years of minimum phase of a sunspot cycle, the surface magnetic fields, with the exception of the polar regions, form a pattern of mixed polarities, i.e., the distribution of the magnetic network elements that, on a spatial scale of at least a supergranule, is a mixture of both polarities in roughly equal portions (Giovanelli, 1982). The polar regions, on the other hand, are covered by predominately unipolar fields that are of opposite polarity in each hemisphere and extend to latitudes of around 50°. As activity levels increase, the mixed-polarity fields at lower latitudes are replaced by active regions and large-scale patterns of unipolar magnetic flux. In concert, the unipolar fields at the poles decrease in size and strength and, during the maximum phase of a cycle, reverse their polarity. As cycle activity declines, the unipolar fields in the polar regions increase in area and strength, while the large-scale unipolar fields are replaced by the time of sunspot minimum with patterns of mixed-polarity fields.

The Butterfly Diagram of the magnetic fields, shown in Fig. 1, illustrates some of their characteristic patterns that occur during a solar cycle. The distribution
Fig. 1. Butterfly Diagram of magnetic flux determined from NSO/KP synoptic rotation maps. Dark shadings represent net magnetic flux that has a negative polarity, and light positive polarity. The times of sunspot minima and maxima are indicated by small and large arrows, respectively. The annual weak variation in the fields near the poles is the result of a varying tilt of the Sun’s axis.

of magnetic flux can be described in terms of two evolving patterns: (1) strong magnetic fields in active regions that early in the cycle appear at mid-latitudes in each hemisphere and progress toward and finally reach the equator approximately 11 years later. The pattern of light and dark shadings of the net flux in the activity belts is a consequence of the strong preference for the leader polarity of active regions to be located equatorward of the follower polarity and the reversal of their polarities with the hemisphere and cycle. (2) weak unipolar bands of both polarities that extend from the activity belts to higher latitudes. The polar magnetic fields reverse when bands of follower polarity reach all the way to the poles.

Kinematic models (e.g. Sheeley et al., 1985, 1989; Wang et al., 1989; Sheeley and Wang, 1994) suggest that the large-scale unipolar patterns and the reversal of the polar fields result from the emergence of magnetic flux in active regions, which then disperses over the solar surface by the processes of diffusion or random walk and by poleward meridional flows. These models closely reproduce the observed distribution of surface magnetic fields and their variation during an activity cycle. The interpretation that active regions are the sole source of the large-scale unipolar magnetic flux patterns is being questioned by several researchers, who suggest that such patterns form in situ from a clustering and preferential alignment of the magnetic poles of many small-scale emerging bipolar regions (Stenflo, 1992; Snodgrass...
and Wilson, 1993) or emerge as a spatially-extended, weak bipolar pattern (Wilson and McIntosh, 1991; McIntosh, 1992).

This paper focuses on the pattern and rate of the emergence of magnetic flux and addresses the question of the source and evolution of large-scale magnetic field pattern. In Section 2, the amount of magnetic flux that emerges in the form of bipolar magnetic regions is estimated and compared with the magnetic flux observed on the solar surface. In Section 3, the evolution of the large-scale magnetic field pattern is discussed in relation to the emergence pattern of active regions and activity nests. The contribution of ephemeral regions to this pattern also is considered. Observations of the possible onset of cycle 23 are presented in Section 4. The conclusions and questions arising from the analyses for this paper are discussed in Section 5.

2. The Magnitude and Rate of Magnetic Flux Emergence

The variation of the magnetic fields, observed in the photosphere during the 19 years of National Solar Observatory/Kitt Peak (NSO/KP) observations, is shown in Fig. 2. The total magnetic flux, i.e. the sum of the absolute value of the positive and negative fields, was determined for each Carrington rotation using the NSO/KP synoptic magnetic maps from 1975 to late 1993. These maps are constructed from the daily, high-resolution, full-disk magnetograms for each Carrington rotation (see, e.g., Harvey, 1994) and represent the magnetic fields when observed near central meridian. Their spatial resolution, 1° in longitude and 0.011 in sine latitude, is reduced from the full-disk data.

These data, which cover cycle 21 and most of cycle 22, show strong differences in the variation in the total magnetic flux between these two cycles. In cycle 22, the total magnetic flux increases more rapidly and reaches higher levels than in cycle 21. For both cycles, the total magnetic flux peaks late in their maximum phase. From the minimum to maximum phases, the total magnetic flux increases on average by a factor of 3.5 in cycle 21 and by 4.0 in cycle 22. For cycle 21, Schrijver and Harvey (1994) estimated the total magnetic flux that emerges in the ensemble of bipolar active regions larger than 520 Mm², the smallest active regions that would be detected in the NSO/KP synoptic maps. Their estimate is based on the following findings: (1) The size distribution of bipolar active regions larger than 520 Mm² roughly follows a power law (Harvey and Zwaan, 1993). (2) The shape of the size distribution of active regions is invariant over the solar cycle, scaling only with the level of activity. For cycle 21, this scaling factor is 8.3 (Harvey and Zwaan, 1993). (3) The relation between the area and magnetic flux in active region plages (excluding sunspots) is linear (Schrijver and Harvey, 1994). (4) The shape of the size distribution of sunspots does not vary with cycle phase (Bogdan et al., 1988).

If the magnetic flux within sunspots of a given size does not vary with the cycle, then the total magnetic flux that emerges as bipolar active regions increases by a factor of 8.3. This value is much larger than the factor of 3.5 increase in the observed total magnetic flux from minimum to maximum in cycle 21 (Fig. 2), indicating that the amount of magnetic flux present on the Sun is not linearly related to the
amount of magnetic flux that actually emerges. Rather, it is a result of the combined processes of magnetic flux emergence, disappearance, and the rates at which these processes take place.

Schrijver and Harvey (1994) find that magnetic flux emerges in the ensemble of active regions at an average rate of $2 \times 10^{21} \text{ Mx/day}$, ranging from $0.7 \times 10^{21} \text{ Mx/day}$ and $6.2 \times 10^{21} \text{ Mx/day}$ during minimum and maximum phases of cycle 21, respectively.

3. The Large-Scale Unipolar Pattern of Magnetic Flux

In Fig. 3, the total magnetic flux on the surface is separated by a threshold of 25 Gauss into active regions and the network elements outside of active regions, the 'quiet' component (Harvey, 1992). A comparison of these two components in Fig. 3 shows that both components vary in phase during a solar cycle, though with substantially different amplitudes. In active regions, the total magnetic flux increases with region area and by a factor of more than 15 from cycle minimum to maximum. The cycle amplitude is no more than a factor of 2 for the quiet component, which is comprised of decaying active region fields and weak network elements. The in-phase variation of the quiet and active region magnetic field components on a time scale of a solar cycle and, on a short time scale, the observed increases in the quiet magnetic network flux that lag pulses of activity in the active component by one to two rotations is consistent with the view that variations in the quiet magnetic fields result from the decay and dispersal of active region magnetic fields.

The origin of the unipolar bands of both polarities in the active region belts is
also suggested in the Butterfly Diagram of the net magnetic flux in Fig. 1. Using finer time resolution than one solar rotation (the width of each column in Fig. 1), the source of these unipolar bands can be traced to the emergence of individual large active regions and activity nests. One such example, shown in Fig. 4, is the emergences of active regions and the subsequent development of the large-scale magnetic pattern that led eventually to the reversal of the north polar fields during cycle 22. The evolution of this magnetic pattern was followed in detail using both the NSO/KP synoptic magnetic maps and full-disk daily magnetograms over an interval of 21 solar rotations from April 1989 to September 1990.

The formation and development of a large-scale pattern of positive polarity (white in the maps) on the poleward side of the active region belt began during Rotation 1813 (one rotation before the first synoptic map in Fig. 4a) at Carrington Longitude 235° and latitude N30°. During the next several rotations, the magnetic flux from this active region expanded in area to higher latitudes and, because of differential rotation, eastward of the Carrington longitude at which the fields originated. Four rotations later (1817), this positive magnetic flux pattern joins another expanding area of positive unipolar fields that developed around longitude 110° over a period of several rotations beginning in Rotation 1816. Over the next 17 rotations, this process is repeated several times as this positive-polarity, large-scale pattern is augmented by a series of subsequent active region emergences, each contributing additional positive polarity flux to the pattern. By rotation 1831 (Fig. 4c), this pattern completely encircles the negative-polarity polar fields. During the emergence and decay of these active regions, their negative-polarity leaders also expand poleward, but for this
Fig. 4. Synoptic magnetic maps of the northern hemisphere covering an interval from April 1989 to September 1990. Lighter and darker shadings represents positive and negative polarity, respectively. Note the E-W orientation of the maps places the follower polarity in active regions to the right of their leader. The starting date, at the upper left of each map, of each Carrington rotation (in parenthesis) is indicated by a vertical bar.
Fig. 4. continued.
Fig. 4. continued.
interval all of this flux cancels with nearby opposite polarity flux. The existing, sporadically augmented, large-scale pattern of positive fields moves toward the pole and by late 1990 or early 1991 reverses the sign of the polar fields.

The evolution of decaying active region flux into large-scale patterns of unipolar, seen clearer as a movie, is a common feature in the redistribution and eventual removal of magnetic flux from the photosphere. The persistence of individual large-scale patterns of unipolar field depends on the spatial distribution and timing of active region emergences and the interactions with neighboring patches of unipolar fields, any one or all of which can either disrupt or augment an existing pattern. These results are in agreement with the strong preferential orientation of large active regions and the lack thereof for smaller regions, together with the power-law like behavior of the size distribution of active regions.

Fig. 5. Butterfly diagram of sunspot regions of cycles 20 (+), 21 (+) and 22 (+) and annual averages of the polarity inversion (heavy curves, solid and dashed for successive cycles). The latitude ranges of ephemeral regions with a preferential orientation reversed from lower-latitude bipolar regions are indicated by rectangular boxes and of small Ca II plage regions by vertical bars. • - the times of the He I 10830 polar coronal holes disappearance; x - completion of the reversal of the polar fields; o - the polar coronal holes reform.
3.1. POLAR CROWN MAGNETIC INVERSION LINE

The reversal of the magnetic fields at the poles takes place over a period of several years as a succession of large-scale bands of unipolar fields move poleward and cancel with the opposite polarity flux in the polar regions (Fig. 1). The latitude of the polarity inversion ($H_t = 0$) at the boundary of the polar fields is a relatively sensitive indicator of the progress of the polar reversal. Previous investigations use polar crown filaments observed in Hα to map this polarity inversion (e.g., Webb, Davis, and McIntosh, 1984; Makarov and Sivaraman, 1989). Known for some times is poleward shift of the polar crown filaments (the so-called rush to the poles) during the rise of a cycle and the disappearance of these filaments once the reversal of the polar fields is completed.

For this study, the latitude of the magnetic polarity inversion closest to the poles in each hemisphere was measured on the NSO/KP daily, full-disk magnetograms at intervals of approximately 7 days in the data. These measurements were checked against the location of the corresponding filament or filament channel. Yearly averages of the latitude of the polarity inversion at the boundary of the polar fields are shown in Fig. 5 in relation to the butterfly diagram of sunspot regions. Also indicated are the range of latitudes over which ephemeral regions are preferentially oriented oppositely to the active and ephemeral regions at lower latitudes. These ephemeral regions are the first bipolar regions of the next sunspot cycle.

The shift in the location of the polemost polarity inversion toward the poles begins with the emergence of sunspot regions during sunspot minimum, as found by Webb, Davis, and McIntosh (1984) and Makarov and Sivaraman (1989), not with the appearance of the first magnetic bipoles of an activity cycle almost three years prior to sunspot minimum. This timing is consistent with the view that active regions play a dominant role in the formation of the large-scale pattern of magnetic flux, which over a several-year period finally results in the reversal of the polar fields and the subsequent and continuing buildup of the polar fields until the next sunspot minimum.

3.2. AXIAL DIPOLE MOMENT OF EPHEMERAL REGIONS

The observed distribution and cycle variation of surface magnetic fields is closely reproduced by kinematic models (Sheeley et al., 1985, 1989; Wang et al., 1989), in which the large-scale unipolar patterns and reversal of the polar fields result from the emergence of active regions, which then are dispersed over the solar surface by the processes of diffusion or random walk and by poleward meridional motions. It has alternatively been suggested that the large-scale pattern emerges in situ as a large-scale, weak magnetic bipole (Wilson and McIntosh, 1991; McIntosh, 1992) or forms from a clustering of numerous small, emerging bipolar regions with a net preferential orientation (e.g. Stenflo, 1992; Snodgrass, 1992; Snodgrass and Wilson, 1993).

To address this later suggestion, the net contribution of ephemeral regions to the Sun's axial dipole moment was estimated using a sample of 9186 regions identified at specific intervals during cycle 21 (Harvey, 1993). Ephemeral regions have lifetimes of several hours and areas of less than 373 Mm$^2$ (an arbitrarily defined upper limit on their size). These small-scale regions emerge at all latitudes, though they occur with
higher frequency in the active region belts. Their numbers vary nearly in phase with the solar cycle, with a minimum occurring about one year before sunspot minimum.

The contribution of the axial dipole moment, $D_{ER}$, of ephemeral regions to the Sun's axial dipole was estimated using the following expression (from Wang and Sheeley, 1991, Eq. 4)

$$D_{ER} = \frac{3}{4\pi R_\odot^2} \Phi \Delta \theta \sin \theta,$$

where $\Phi$ is the magnetic flux in one pole of a dipole (in Mx), $\Delta \theta$ the separation of the region's poles in latitude (in radians), and $\theta$ the co-latitude measured from the North Pole. The sign of the axial dipole moment is defined by the polarity of the region's pole closest to the Sun's North Pole. $\Delta \theta$ was estimated using an average separation of an ephemeral region's poles (in this case 13,500 km) and on the region's orientation crudely determined within 45° bins (see Harvey, 1993, Chapter 5). Although the orientation bins are wide, the location of the leader polarity relative to the position of the follower and, therefore, the sign of the axial dipole moment, is well determined.

In Fig. 6, the net contribution of ephemeral regions to the Sun's axial dipole moment, summed in each year, is shown in relation to that determined for a sample of 978 active regions observed during the same time intervals as the ephemeral regions (Harvey, 1993; Harvey and Zwaan, 1993), as well as with that determined by Wang and Sheeley (1991) for a sample of about 2700 active regions. Both of the active region samples were identified on the NSO/KP daily full-disk magnetograms.
In the data set used by Harvey and by Harvey and Zwaan, active regions, identified only during selected intervals, are corrected for several effects that reduce their visibility (as are the ephemeral regions). Wang and Sheeley included all active regions observed to emerge between 1976 and 1986. When scaled to a common time interval (i.e. one year), the net axial dipole moment of the two active region sample is comparable. It should be noted that summing the net contributions of magnetic bipolar regions to the axial dipole moment of the Sun over one year assumes that the magnetic flux that emerges in bipolar regions remains on the surface for that period.

As shown in Fig. 6, the net axial dipole moment of ephemeral regions is of the opposite sign to the net contribution of active regions, and its amplitude smaller by a factor of 6, assuming an average lifetime of ephemeral regions of 8 hours. (Ephemeral regions are identified on magnetograms taken about 24 hours apart, a longer interval than the average lifetime of these bipolar regions. The number of ephemeral regions that emerge per day can be estimated by multiplying by the ratio of 24 hours and their average lifetime.)

During cycle 21, the Sun's axial dipole moment varied from a maximum positive value of +3.45 Gauss in 1976 to a maximum negative value of -4.0 Gauss in the following sunspot minimum in 1986, reversing its sign in late 1979 to 1980 (Wang and Sheeley, 1991). The respective contributions, shown in Fig. 6, indicate that it is not ephemeral regions, but rather active regions that are the cause of the reversal of the Sun's axial dipole moment.
SOLAR MAGNETIC CYCLE

The contribution of bipolar active regions to the Sun's axial dipole moment, summed over the 12-year interval covering cycle 21, is shown in Fig. 7 as a function of region size, determined from the sample of 9186 ephemeral regions and 978 bipolar active regions (Harvey, 1993). The bipolar active regions that contribute the largest share of the negative axial dipole moment are those regions with areas greater than 4500 Mm$^2$. This result agrees with Wang and Sheeley (1991), who find that the larger active regions in their sample are the largest and most significant contributors to the reversal of the axial dipole moment of the Sun.

The distribution in latitude of the axial dipole moment of ephemeral regions is compared in Fig. 8 with the Butterfly Diagram of the net magnetic flux determined (from the NSO/KP synoptic magnetic maps) in the same time intervals during which ephemeral regions were identified. With time and latitude, the sign of the contribution of ephemeral regions to the Sun's axial dipole moment shows an almost random pattern that appears unrelated to either the polarity or pattern of observed magnetic flux. Based on these data, it is difficult to see how the emergence of ephemeral regions could contribute significantly to the formation of the large-scale patterns of unipolar fields, since in order to do so, they must show a net alignment in their orientations spatially organized on a large-scale. Figure 8 suggests that the component of the orientation of ephemeral regions in the north-south direction is not systematically organized on time scales of a few rotations or on spatial scales comparable to the large-scale unipolar field patterns.

4. Evidence of New Cycle 23 Activity?

The first bipolar regions of an activity cycle are small active regions and ephemeral regions (Martin and Harvey, 1979; Harvey, 1992). As shown in Fig. 9, this pattern of bipolar emergence forms an extended butterfly diagram of active regions that begins at high latitudes about two years after the reversal of the polar fields and systematically and smoothly progresses into the later developing pattern of sunspot regions.

Since 1990, sequences of NSO/KP magnetograms were made each year during the spring and fall in an effort to look for the onset of cycle 23. A preliminary analysis of these data by K. Harvey and S. Martin suggests that high-latitude ephemeral regions and small active regions began showing a preferential orientation reversed from lower-latitude regions in the spring of 1993. The timing of the development of this pattern is similar to the two-year interval between the reversal of the polar fields and the first appearance of preferentially reverse-orientation ephemeral regions that has been observed in the two previous cycles (Martin and Harvey, 1979; Harvey, 1992). It is suggested that these regions may be the first evidence of the onset of cycle 23. Before a conclusive statement can be made, however, further observations are needed to verify that emergence of preferentially reverse-orientation ephemeral and small active regions continues and forms the beginnings of the extended butterfly diagram of the magnetic bipolar regions of cycle 23.
5. Conclusions and Discussion

The spatial and temporal variations of the magnetic fields observed in the photosphere are a consequence of the emergence of magnetic flux, the disappearance or cancellation of magnetic flux, and the rates at which these processes take place. Most of the magnetic flux that emerges within an active region disappears in situ,
with only a small fraction expanding away from its emergence site as the region decays. Increases in the magnetic flux present on the Sun result from a higher rate of magnetic flux emergence than flux disappearance, while decreases indicate a higher rate of magnetic flux disappearance than flux emergence.

The rate of magnetic flux emergence in the ensemble of active regions in cycle 21 (on average of $2 \times 10^{21} \text{ Mx/day}$) is a factor of 10 less than in ephemeral regions ($2 \times 10^{22} \text{ Mx/day}$). In terms of the magnitude of magnetic flux emergence, ephemeral regions are not negligible. Their short lifetimes and small size scale, however, suggest that despite their large numbers, ephemeral regions have only a transient effect on the magnetic flux present on the solar surface. Per day, the total amount of flux emerging in ephemeral regions is at most 5% of the magnetic flux in the quiet Sun, and within hours it disappears from the surface.

Ephemeral regions emerge over a much wider latitude range than active regions and with only a slight preferential orientation that is proper for their hemisphere and cycle, as defined by the Hale-Nicholson Polarity Law for active regions. For the 9186 ephemeral regions, identified during selected intervals throughout cycle 21, there is a tendency for the positive-polarity pole to lie north of the negative-polarity pole. This slight tendency results in a net positive contribution of ephemeral regions to the Sun's axial dipole moment, opposite to the negative axial dipole moment contributed by active regions during cycle 21. About one-third of the total axial dipole moment contribution of ephemeral regions comes from those regions emerging poleward of 35° latitude.

Based on these data, the contribution of ephemeral regions to the Sun's axial dipole moment appears to counteract that of active regions, although only by a relatively small amount. This amplitude of the axial dipole moment contribution of ephemeral regions, however, is based on fairly crude measurements of the orientation and on an average size and total magnetic flux of ephemeral regions. More accurate measurements of the orientation, magnetic flux, the separation of the poles, and longitude distribution of ephemeral regions are needed to better assess the effect of these regions on the Sun's axial dipole moment and on the evolution or maintenance of the large-scale pattern of unipolar magnetic fields. From this study, however, there is no evidence to suggest that ephemeral regions are effective in forming the large-scale patterns of unipolar magnetic fields. As argued by Howard (1992), a careful detailed analysis of the time-varying distribution of surface magnetic fields shows that all large-scale unipolar patterns can be traced to the emergence of magnetic flux in active regions and activity nests.

Over much of the cycle, ephemeral regions emerge within large-scale patterns of unipolar fields. Ephemeral regions have a lifetime of typically only a few hours, ending when one of their poles cancels with nearby opposite polarity network elements. Hence, the net magnetic flux contribution of an ephemeral region to the unipolar fields is zero. The process of the emergence of an ephemeral region and the cancellation of the polarity opposite to the surrounding network elements, however, provides a mechanism to redistribute magnetic flux on the surface, but only over distances of less than the diameter of a supergranule. This mechanism, first proposed by Marsh (1978), was considered by Wang and Sheeley (1991) in their model of flux transport, though they assumed that the orientations of ephemeral regions are random. In their
Fig. 9. Butterfly diagram of sunspot regions. Regions with sunspot area >100 millionths of a hemisphere are shown alternately for successive cycles by + and o. The rectangular boxes indicate the latitude range of ephemeral regions with a preferential orientation reversed from lower-latitude regions, and the light vertical bars the of small Ca II plage regions. The last sunspot region belonging to a cycle is shown by a o and the times of sunspot minima (m) by arrows.
SOLAR MAGNETIC CYCLE

model, ephemeral regions added a 'noisy, coarse-grained' structure to the large-scale fields, but they did not cause or participate in the poleward migration of magnetic flux. It would be instructive to repeat these calculations using the observed slight preferential orientation and latitude distribution of ephemeral regions.

Acknowledgements. My thanks to C. Zwaan, C. Schrijver, S. D'Silva, R. Howard, J. Harvey, and J. Stenflo for the many helpful discussions regarding this ongoing research. This work was supported under NASA Grant NASW-4721 and by NATO Collaborative Research Grant 910997 and was done while a visitor at the National Solar Observatory (National Optical Astronomy Observatories, operated by the Association of Universities for Research in Astronomy, Inc. (AURA), in cooperative agreement with the National Science Foundation). The NSO/Kitt Peak data used here are produced cooperatively by NSF/NOAO, NASA/GSFC, and NOAA/SEL.

References
