THE REVERSAL OF THE SOLAR POLAR MAGNETIC FIELDS

I. *The Surface Transport of Magnetic Flux*

P. R. WILSON

Department of Applied Mathematics, University of Sydney, Sydney, 2006 Australia

P. S. McINTOSH

Space Environment Laboratory, NOAA, 325 Broadway, Boulder, CO 80303, U.S.A.

and

H. B. SNODGRASS

Physics Dept., Lewis and Clark College, Portland, OR 97219, U.S.A.

(Received 21 November, 1989)

Abstract. Some theoretical difficulties confronting the current model of the polar magnetic reversal by cancellation with the flux remnants of decaying active regions are discussed. It is shown that the flux transport equation does not adequately describe the essential physical consequences of the transport of large-scale fields, linked to deep subsurface toroids, over distances comparable with the solar radius. The possibility that subsurface reconnections may release these fields to form U-loops is discussed but it is shown that, in this event, the loops will quickly rise to the surface. Mechanisms whereby the flux may escape through the surface are considered.

1. Introduction

The first observations of the reversal of the Sun's weak polar magnetic fields were obtained during cycle 19. At the beginning of the cycle, the polar polarity of a given hemisphere corresponded to that of the leader spots of that hemisphere but, in 1957, the magnetic polarity of the south pole was observed to reverse and, eighteen months later, that of the north also changed. Thus at the start of cycle 20, the polar polarities in each hemisphere again corresponded to those of the leader spots. Similar reversals also occurred shortly after maximum during cycles 20 and 21.

Although the polar fields are weak, even at minimum, their reversal seems to 'set the stage' for the new cycle. Shortly after reversal, the large-scale field configuration changes qualitatively and the first phenomena of the extended cycle (Wilson *et al.*, 1988) appear at high latitudes, showing polarity patterns consistent with those expected for sunspot pairs of the following cycle and with the (now reversed) polar fields. Thus the polar fields appear to be an integral part of the activity cycle and, since poloidal fields are an essential component of dynamo models, it is important to understand the part played by the polar fields and their reversals.

The current cycle is building rapidly towards maximum and, in order to complement observational studies of the polar polarity reversals, this series of papers will discuss existing models of the polar field reversal, explore some of their observational conse-

Solar Physics **127**: 1–9, 1990.

quences, and suggest a new model. In the first paper, we consider the flux cancellation model, in which the reversal is due to the preferential equatorward diffusion and trans-equatorial cancellation of the remnants of leader flux from decaying active regions and the poleward motion of the remaining flux, now with an excess of follower flux, which cancels with the existing polar flux. We discuss here some theoretical difficulties associated with this model.

2. The Flux Cancellation Model

An important feature of the model is the concept that, as the active region fields decay, the diffusive effect of the supergranule motions distributes the leader and follower flux over a considerable area, and gives rise to the large-scale fields. Because of the initial tilt of the sunspot axes (leader equatorward), it is assumed that the regions of leader polarity diffuse preferentially towards the equator where they cancel with leader flux from the opposite hemisphere. The remaining large-scale regions, among which those of follower flux are now in the majority, diffuse polewards, assisted now by a postulated poleward meridional flow. There they contribute to the reversal (by cancellation with existing flux) and subsequent build-up of new cycle polarity flux there. Leighton (1964) first proposed the random walk model, using supergranules as the stepping scales and DeVore *et al.* (1985) described this migration in terms of a diffusion model, with diffusivity $300\ \mathrm{km}^2\ \mathrm{s}^{-1}$, augmented by variable meridional poleward flows of $\sim 10\ \mathrm{m\ s}^{-1}$. This model has been further developed by Sheeley, Nash, and Wang (1987) and has had some success in reproducing some of the patterns of evolution of the large-scale field. Furthermore, the correlation-tracking of granules indicates that the transport of magnetic flux is well correlated with horizontal photospheric flows.

Despite these successes, models based on the flux transport equation (DeVore *et al.*, 1985) are subject to the criticism that they consider only the radial component of the magnetic field, treating it as a *scalar* quantity. Indeed, in Leighton's paper, he defines only the 'number surface density, *n*, of points at which lines of force enter the Sun' and derives his fundamental equation, on which his model depends (Equation (12)), in terms of *n*. DeVore, Sheeley, and Boris (1984), in an Appendix, attempt a more detailed derivation of an equivalent equation, in which they write both the magnetic and velocity fields in terms of an ensemble averaged term and a fluctuating term, i.e.,

$$\mathbf{B} = \langle \mathbf{B} \rangle + \delta \mathbf{B}, \qquad \mathbf{v} = \langle \mathbf{v} \rangle + \mathbf{v}. \tag{1}$$

Thus $\langle \mathbf{B} \rangle$ represents the large-scale magnetic fields and $\langle \mathbf{v} \rangle$ the global rotation and meridional circulation. Taking the ensemble average of the induction equation, they obtain

$$\frac{\partial \langle \mathbf{B} \rangle}{\partial t} - \nabla \times (\langle \mathbf{v} \rangle \times \langle \mathbf{B} \rangle) = \nabla \times (\langle \delta \mathbf{v} \times \delta \mathbf{B} \rangle). \tag{2}$$

They then consider only the radial component of Equation (2) and obtain

$$\frac{\partial \langle B \rangle_r}{\partial t} - \nabla \cdot (\langle B \rangle_r \langle \mathbf{v} \rangle_s) = \mathbf{r} \cdot \nabla \times (\langle \delta \mathbf{v} \times \delta \mathbf{B} \rangle), \tag{3}$$

where $\langle B \rangle_r$ is the radial component of the large-scale field and $\langle \mathbf{v} \rangle_s$ is the vector surface component of the large-scale velocity field. The right-hand side of Equation (3) is then evaluated by making the rather surprising assumption that $\langle \mathbf{v} \rangle = 0$ and this, after some algebra, yields their flux transport equation,

$$\frac{\partial \langle B \rangle_r}{\partial t} - \nabla \cdot (\langle B \rangle_r \langle \mathbf{v} \rangle_s) = \kappa \nabla_s^2 \langle B \rangle_r . \tag{4}$$

Two comments are appropriate here. First the assumption that $\langle \mathbf{v} \rangle = 0$, which is made in order to evaluate the right-hand side of Equation (3), necessarily entails *all* the components of $\langle \mathbf{v} \rangle$ are zero, in particular, the surface component, $\langle \mathbf{v} \rangle_s$. Thus Equation (4) reduces to the diffusion equation,

$$\frac{\partial \langle B \rangle_r}{\partial t} = \kappa \nabla_s^2 \langle B \rangle_r , \tag{5}$$

which is unfortunate if a meridional flow velocity is included in the modelling process.

Secondly the neglect of all but the radial component of Equation (2) is justified on the grounds that the large-scale field is predominantly radial. While this may be true near the surface, the neglect of the other components of Equation (2) is justified only if $\langle \mathbf{B} \rangle$ remains predominantly radial over distances along the field lines which are comparable with the transverse scale of the field; otherwise the vector nature of the large-scale field and, in particular, the effects of curvature, are ignored.

However, the bipolar magnetic fields which appear at the solar surface in active regioins are assumed to arise from the emergence of a loop or stitch of flux, sometimes characterized as an 'Ω-loop', in a postulated sub-surface toroid anchored at the base of the convection zone and, in situations where surface flux is transported across distances comparable with the solar radius, the assumption that the fieldlines remain predominantly radial is open to considerable doubt.

In such a case, the neglect of the non-radial components of Equation (2) effectively treats the surface field elements as if they were 'corks' carried along by an eddying stream, ignoring the consequences of the continuity of magnetic field lines, entailed by Maxwell's law ($\nabla \cdot B = 0$). While the surface fields remain connected to this toroid, a more appropriate analogy is with a marine 'buoy' anchored to the bottom of a tidal flow by an elastic cord. When the buoyant surface elements of these fields are displaced through any significant distance by the drag of the meridional flow, the induced curvature of the field lines gives rise to a net transverse component of the magnetic tension which, acting on an element of the flux tube, may balance the drag force of the flow.

The magnitude of this effect may be estimated by considering an element of the flux tube of radius r, length δh and field B. Neglecting diffusion and considering only the effect of a transverse flow v, the drag force δF is

$$\delta F = C \rho v^2 r \delta h , \tag{6}$$

where ρ is the plasma density and C is the drag coefficient of order unity. If, as a result of this drag, the element is inclined to the vertical at angle θ at its upper end and at $\theta + \delta\theta$ at the lower end, the component of the magnetic tension opposing the drag, δT_m, is

$$\delta T_m = \tfrac{1}{4} r^2 B^2 \cos\theta\, \delta\theta . \tag{7}$$

Thus the drag force is balanced when

$$\delta\theta = \frac{4 C \rho v^2 \delta h}{r B^2 \cos\theta} . \tag{8}$$

For a flux tube with field of order 10^3 G and of radius ~ 1000 km embedded in a plasma of mean density 4×10^{-4} g cm^{-3} with a transverse flow of 10 m s^{-1}, and taking C as unity, a distortion, $\delta\theta$, of only 0.016 in an element of length 10^4 km would be sufficient to permit the nett magnetic tension to balance the drag force. Thus it is hard to see how meridional flows of order 10 m s^{-1} could significantly deform the subsurface flux tubes associated with the surface fields, far less transport them either to the polar regions in order to contribute to the polar reversals or to take part in trans-equatorial cancellations. Furthermore, even if significant deformation occurs, the net transverse component of the magnetic tension would increase until it balanced the drag force of the flow, just as the anchor rope of a floating buoy is deformed. We conclude that, whatever its success in replicating the apparent surface displacements, the flux transport equation does not fully describe the physical processes associated with the surface transport of the large-scale fields.

3. Sub-Surface Reconnections

Nevertheless, the observed surface fields *appear* to undergo significant displacements and it may be that the effects of field line distortion and magnetic tension may be overcome by reconnections between adjacent elements of the distorted field. Certainly, field line reconnections in the corona were observed by Skylab and are also seen in chromospheric Hα observations, so there is no reason why they should not occur below the surface and thus release the surface fields from the constraints of the deep magnetic toroids.

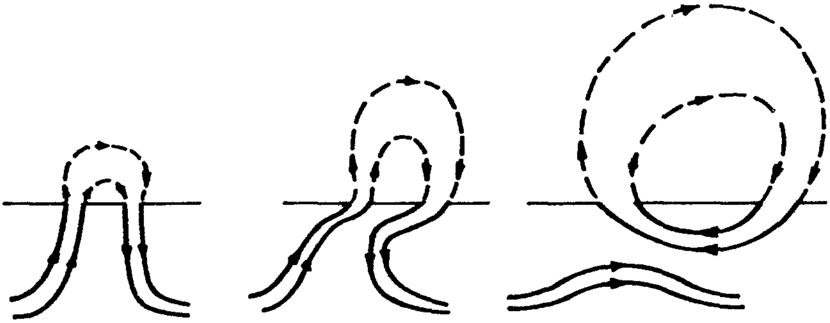

Fig. 1. The formation of a U-loop by the 'repair' of a sheared Ω-loop is illustrated.

A possible example of reconnection, leading to the formation of a 'U-loop' is illustrated in Figure 1, the process being consistent with that proposed by Spruit, Title, and van Ballegooijen (1987) which they call the 'repair' of the Ω-loop. However, when the U-loop forms, it must necessarily be buoyant since, under isothermal conditions, the plasma density in both the base and the sides of the 'U' must be less than the ambient density in order to accomodate the magnetic pressure within the 'U'. Thus initially it must rise through the region. As it rises, however, plasma drains down the sides to the base and, assuming that the plasma cannot cross the field lines, it remains trapped within the 'U'. Although the density within the 'U' thereby increases, the dimensions of the horizontal section of the tube must expand because, in order to maintain pressure balance under isothermal conditions, the density within the tube must remain slightly less than that outside (see Wilson, 1989). Thus, the 'U' is always buoyant and must rise towards the surface in a configuration which is illustrated in Figure 2(a).

By analogy with the buoyancy of a bubble of gas, the speed of the rising flux might be comparable with the Alfvén speed and thus the rise time from a depth of (say) 10 000 km would be of the order of a few weeks, i.e., less than 1% of the time required for the field to drift or random-walk to the poles from latitudes of $\sim 30°$, assisted by meridional flows of $\sim 10 \text{ m s}^{-1}$.

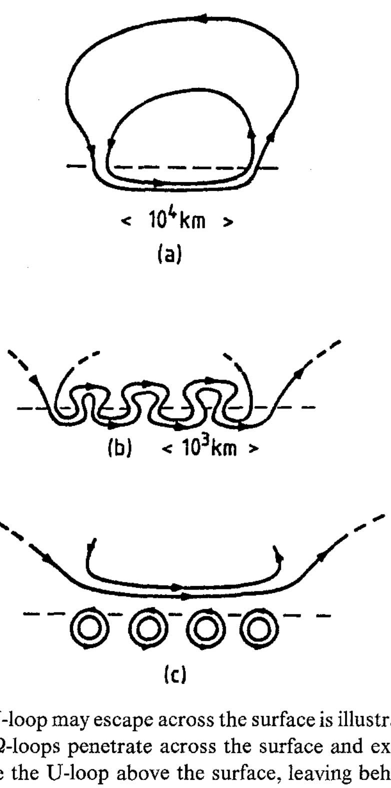

Fig. 2. The process by which a U-loop may escape across the surface is illustrated. (a) The U-loop has risen to the surface region. (b) Small Ω-loops penetrate across the surface and expand as plasma drains down the 'U'. (c) Reconnections release the U-loop above the surface, leaving behind small closed loops which decay rapidly.

4. The 'Escape' of the U-Loop

It is not immediately clear, however, that the U-loop can now escape. Parker (1984) has argued that purely toroidal fields which have been elevated to the surface of a star may be unable to escape through the surface because of the difficulty of disengaging the field from the highly conducting gas in which it is embedded. He suggests that the appropriate boundary condition is $\partial B_\phi / \partial r = 0$, which represents a closed boundary, rather than an open boundary, through which flux may readily escape, described by $B_\phi = 0$.

Parker describes a possible 'escape scenario' in which neutral point reconnections between adjacent Ω loops above the surface can effect the release of toroidal field lines. However, he points out that, even at sunspot maximum, the gaps in longitude between adjacent bipolar magnetic regions are so wide that the reconnection process would be severely limited.

On the other hand, the topology of the U-loop structures resulting from the sub-surface reconnections considered here is rather different for, although the base of the 'U' corresponds to an element of toroidal field, the sides of the 'U' break through the surface and the complete structure is closed above the surface (see Figure 2(a)). As Parker has shown, the field is in a marginally unstable state just below the surface and would be susceptible to the Kelvin–Helmholtz instability. Thus it is vulnerable to small-scale temperature or velocity perturbations, and small elements of the now dispersed and weaker field lines may bend upwards. As such kinks or small loops develop, plasma drains down the field lines, making the crests more buoyant so that they may cross the surface, forming small Ω-loops, as illustrated in idealized form in Figure 2(b).

This configuration is similar to that envisaged by Parker but on a 'granule sized' scale or smaller. Furthermore, since these loops may be more tightly packed (on this scale), reconnections may proceed as illustrated in Figure 2(b). After a series of such reconnection, the U-loop would be located entirely above the surface and is free to escape, leaving behind a series of closed loops (Figure 2(c)) whose decay is discussed below.

5. Ephemeral Active Regions

While the configurations illustrated in Figure 2 are somewhat idealized, it might be expected that, if much of the decaying flux from active regions escapes in this way (although in a less idealized manner), the process should be observable. We suggest here that it is consistent with the observed pattern of emergence and decay of ephemeral active regions (ERs).

ERs are small magnetic bipoles of mean total flux $\sim 10^{19}$ Mx, major axes $\sim 10\,000$ km and typical lifetimes of less than one day. They have been studied extensively by Martin and Harvey (1979) who have argued that they form part of the general spectrum of active regions. However, their number distribution against size does not fit smoothly with the overall active region distribution, ERs being considerably more numerous than would be expected at this end of the AR distribution. They are also more widely distributed over the solar surface. The total surface flux in the form of ERs at any given time is $\sim 10^{22}$–10^{23} Mx, which is comparable with that of a large active region or complex and so, with lifetimes of less than a day, the ER population represents a significant rate of appearance and disappearance of surface magnetic flux.

An ER first emerges as a bipolar pair of flux knots in the neighbourhood of elements of the large-scale field or the network. As the knots separate, one or other approaches a field element of opposite polarity, where cancellation takes place. We postulate that this process of emergence and cancellation is related to the escape of the flux from decaying active regions trapped just below the surface and, in Figure 3, the posible reconnections related to the emergence and decay of one ER are illustrated diagramatically.

Each reconnection releases a part of the trapped U-loop, leaving behind a single closed loop and a smaller trapped U-loop, and, as other Ω-loops emerge from the trapped U-loop in the form of ERs, further reconnections continue the process of releasing closed flux loops and decreasing of the size of the trapped U-loops until their radius of curvature approaches that of the closed loops (Figure 3(c)).

Thus the emergence of ERs may well correspond to a more random version of the idealized process illustrated in Figure 2(b) and, if the emergence and disappearance of ERs does indeed represent a rate of flux loss of $\sim 10^{22}$ Mx day^{-1}, it would be more than sufficient to eliminate all the active region flux present on the Sun within a few weeks.

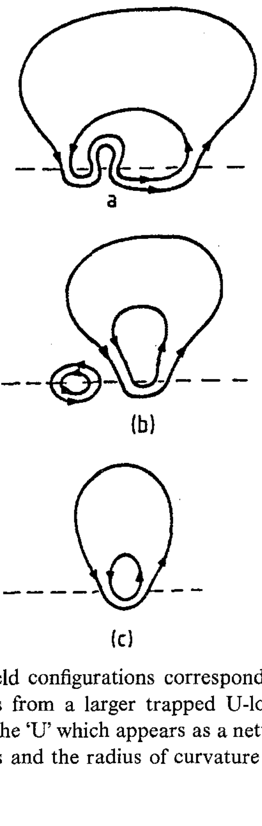

Fig. 3. The postulated sub-surface field configurations corresponding to the emergence of an ER as a bi-polar Ω-loop. (a) The loop emerges from a larger trapped U-loop. (b) As the Ω-loop expands, one component reconnects with one leg of the 'U' which appears as a network element, releasing a closed loop. (c) The closed loop decays and the radius of curvature of the U-loop decreases.

Indeed it would indicate the need for a source of surface flux in addition to that available, in principle, from the decay of active regions. However, even if the escape process were somewhat less efficient, it would seem unlikely that the transverse component of a buoyant U-loop can remain trapped below the surface by the Parker boundary condition for periods sufficient to permit the postulated surface migrations.

6. The Decay of Closed Flux Loops and U-Loops

There remains the question of the small closed flux loops which arise as part of this process. The theoretical decay time for a closed or partly closed loop of flux of radius of curvature R_0 is of order R_0^2/η, where η is the resistive diffusivity. This reaches a maximum of only 10^9 cm^2 s^{-1} at the surface of a star like the Sun (Kovita and Cram, 1983), and is much smaller in the plasma above and below the surface. Thus for U-loops with initial dimension $\sim 10\,000$ km, the ohmic decay time is of the order of 30 years, while for closed loops formed by the processes described above, having dimensions of order $\sim 1\,000$ km, it would still be of the order of several months to a year. However, a closed loop of flux is inherently unstable, since the net magnetic tension acts inwards along the radial direction and, for a loop embedded in an otherwise unrestricted plasma, it is unlikely that a radial pressure gradient can be built up to support the magnetic net tension and it must necessarily collapse inwards.

The time-scale of this collapse is determined by a balance between the component of the magnetic tension, acting inwards, and the viscous drag force. If r is the radius of the cross-section of the loop, R its radius of curvature, and ρ is the density of the surrounding plasma, then, following Equation (6), the drag force on an element subtending an angle $\delta\theta$ and collapsing inwards with speed v is

$$\delta F = C\rho v^2 r R\,\delta\theta\,. \tag{9}$$

Again, following Equation (7), the radial component of the magnetic tension on the element is

$$\delta T_m = \tfrac{1}{4}B^2 r^2\,\delta\theta\,, \tag{10}$$

and these are balanced when

$$v^2 = \frac{B^2 r}{4\rho R C}\,. \tag{11}$$

Thus the decay time, τ, is of order

$$\tau \sim \frac{R}{v} = \frac{2R^{3/2}C^{1/2}\rho^{1/2}}{B r^{1/2}}\,. \tag{12}$$

For a flux tube of cross-sectional radius 1000 km, in a closed loop of curvature 10000 km, field 1000 G embedded in a plasma of density 2×10^{-7} g cm^{-3}, taking C of order one yields $\tau \approx 4 \times 10^3$ s.

Even if the density gradient or other effects were to increase the collapse time by an order of magnitude, it would be at most of order one day and we coclude that the closed loop structures, formed as part of the Ω-loop reconnection process, would be unable to resist the combination of collapse and decay and would be rapidly destroyed.

7. Conclusion

It has been shown that the flux transport equation does not adequately describe the physical processes which must be associated with the surface transport of flux from decaying active regions over large distances. Basic theory requires that this flux should remain confined to the region of emergence by the sub-surface field line tensions. It is posutlated that sub-surface reconnections giving rise to U-loop formation might release the surface fields from the constraints of magnetic tensions but, in this event, it is shown that buoyancy would bring the U-loops to the surface and a mechanism is described whereby the transverse field component can leak through the surface layers in a time too short to enable it to take part in the processes described by the flux transport equation; i.e., the trans-equatorial cancellation of leader flux and the polar field reversals by the net remaining follower flux.

Nevertheless flux apparently related to decaying ARs is observed to last in coherent form for several solar rotations and does appear to migrate polewards. However, according to Howard and LaBonte (1981), this migration is episodic rather than continuous, and so inconsistent with a diffusive process. Thus our understanding of the phenonena is less than complete and it would seem that either the escape mechanism does not apply in the Sun or that some form of flux regeneration and propagation must be associated with decaying active regions.

These questions will be further explored in the next paper in this series.

References

DeVore, R. C., Sheeley, N. R., and Boris, J. P.: 1984, *Solar Phys.* **92**, 1.
DeVore, R. C., Sheeley, N. R., Boris, J. P., Young, T. R., and Harvey, K. L.: 1985, *Australian J. Phys.* **38**, 999.
Howard, R. F. and LaBonte, B. J.: 1981, *Solar Phys.* **75**, 131.
Kovita, P. and Cram, L. E.: 1983, *Solar Phys.* **84**, 45.
Leighton, R. B.: 1964, *Astrophys. J.* **140**, 1547.
Martin, S. F. and Harvey, K. L.: 1979, *Solar Phys.* **64**, 93.
Parker, E. N.: 1984, *Astrophys. J.* **281**, 839.
Sheeley, N. R., Nash, A. G., and Wang, Y-M.: 1987, *Astrophys. J.* **319**, 481.
Spruit, H. C., Title, A. M., and van Ballegooijen, A. A.: 1987, *Solar Phys.* **110**, 115.
Wilson, P. R.: 1989, *Proc. Astron. Soc. Australia* **8**, 32.
Wilson, P. R., Altrock, R. C., Harvey, K. L., Martin, S. F., and Snodgrass, H. B.: 1988, *Nature* **333**, 748.