Life and Death of Solar Active Regions

A. A. van Ballegooijen

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Abstract. Recent models for the formation and decay of active regions are discussed. Large active regions are thought to be Ω loops that emerge from a toroidal field located near the base of the convection zone. After an Ω loop has fully emerged at the solar surface, it continues to evolve under the influence of subsurface convective flows and solar differential rotation. The magnetic helicity of active regions originates below the photosphere, but its spatial distribution in the corona is significantly altered by reconnection processes. Results from 3-D flux transport models of decaying active regions are presented. It is found that such models provide a natural explanation for the occurrence of flares and coronal mass ejections. The models predict that submerging magnetic fields are transported back to the base of the convection zone, leading to the “repair” of the toroidal flux ropes. Interactions between Ω loops are also considered.

1. Introduction

Magnetic flux emerges at the solar surface in the form of bipolar active regions. The properties and emergence patterns of active regions have been studied in great detail by Harvey & Zwaan (1993). Large regions generally have an East-West orientation of magnetic polarities in agreement with the Hale-Nicholson law (Hale et al. 1919; Hale & Nicholson 1925). This suggests that these regions are Ω loops emerging from a toroidal field located near the base of the convection zone (Parker 1955; Babcock 1961; Zwaan 1996). In this paper I review recent work on the formation and decay of Ω loops.

2. Flux Emergence

There is a tendency for new flux to emerge in the vicinity of an existing active region (Bumba & Howard 1965; Gaizauskas et al. 1983; Castenmiller, Zwaan, & van der Zalm 1986; Brouwer & Zwaan 1990; Zwaan & Harvey 1994). This indicates that the underlying toroidal field is already disturbed and is likely to erupt again. Early in the life of a large region, the observed magnetic flux increases by the emergence of several distinct bipoles. The elements of each bipole separate from each other, and like-polarity elements from different bipoles may coalesce into sunspots, producing a complex magnetic structure with multiple interacting bipoles. Initially, a significant fraction of the magnetic flux is contained in sunspots, with the remainder in smaller magnetic elements. The leader polarity flux is usually more concentrated than the follower flux. It is unclear what holds these leader spots together.
The physical processes that produce $\Omega$ loops in the convection zone (CZ) are not well understood. One possibility is that the loops are produced by buoyancy-driven instabilities of the toroidal field (e.g., Caligari, Moreno-Insertis, & Schüssler 1995; Fan 2001, 2004). Due to the stabilizing effects of magnetic tension, the most unstable modes of the toroidal field are those with the lowest azimuthal mode number ($m = 1, 2, 3, \cdots$). Therefore, buoyant instabilities tend to produce wide $\Omega$ loops that extend more than 90° in longitude, much larger than even the largest active regions. This suggests that magnetic buoyancy is not the only force at work in the formation of an $\Omega$ loop; interactions of the magnetic field with deep-seated convective flows may also be important (e.g., Dorch et al. 2001). In any case, once an $\Omega$ loop is produced it will likely be further distorted by convective flows as it rises through the CZ (Abbett et al. 2004). The observed tilt of active regions (Joy’s law) is presumably due to Coriolis forces acting on the rising $\Omega$ loop. Interactions with helical flows can cause the field to become twisted (Longcope et al. 1998).

Due to the stratification of the CZ, the rising flux tubes have a tendency to expand and disperse horizontally. Therefore, several authors have suggested that the magnetic field of an $\Omega$ loop must be twisted together into a rope-like structure, so that the flux bundle rises to the surface as a coherent structure and emerges as one active region (Linton et al. 1999; Fan et al. 1999; Archontis et al. 2004, 2005). Magara (2006) has simulated the emergence of a twisted flux rope into the solar atmosphere and finds qualitative agreement of such simulations with EUV and X-ray observations.

Observations of photospheric vector fields in active regions indicate that the magnetic field can be strongly sheared in the direction along the polarity inversion line (e.g., Pevtsov, Canfield, & Metcalf 1995; Pevtsov, Canfield, & McClymont 1997). X-ray observations show “sigmoids” indicating that the coronal fields above such regions are highly sheared or twisted (Acton et al. 1992; Rust & Kumar 1996; Schmieder et al. 1996; Canfield, Hudson, & McKenzie 1999; Sterling et al. 2000). Therefore, active regions have nonzero magnetic helicity. The shear or twist originates below the photosphere (Leka et al. 1996; van Driel-Gesztelyi et al. 1997; McClymont, Jiao, & Mikić 1997). Pevtsov, Maleev, & Longcope (2003) have shown that it takes about one day for magnetic twist to propagate into the corona from the subsurface layers. Rust (1994) argues that the twist originates in the toroidal field at the base of the CZ. According to this view the toroidal field consists of flux ropes with left (right)-helical twist in the northern (southern) hemisphere.

3. Decay of Active Regions

As an active region evolves, there is a remarkable transition from a coherent $\Omega$ loop moving under the influence of magnetic buoyancy to a passively diffusing, readily deformable field configuration (Schrijver & Title 1999). The sunspots decay and the magnetic flux is dispersed over the solar surface due to the interaction of the $\Omega$ loop with subsurface convective flows and the solar differential rotation. This process is illustrated in Figure 1, which shows the evolution of an active region in the Southern hemisphere over a 3-month period in 2005. As the remnant region disperses, the average magnetic flux density decreases and
opposite polarity fluxes cancel each other at polarity inversion lines (PIL) on the photosphere (Martin, Livi, & Wang 1985). Eventually, the remnant fluxes from neighboring regions cancel or merge, so that the original region can no longer be clearly distinguished.

The subsurface structure of decaying active regions is not well understood. If there is a direct magnetic connection between the vertical field at the solar surface and the horizontal field near the base of the CZ, the magnetic tension in the toroidal tubes will pull the two polarities of an active region apart in the longitudinal direction (van Ballegooijen 1982). The longitudinal drift velocity \( v_d \) depends on the hydrodynamic drag exerted on the vertical part of the flux tube. The predicted velocity can be quite large: for an adiabatic flux tube anchored in a toroidal tube with field strength 25 kG and radius 500 km below the base of the CZ, the predicted velocity \( v_d \sim 60 \text{ m/s} \). Observations of decaying active regions do not show such systematic drifts (upper limit is \( \sim 10 \text{ m/s} \)). This implies that either (1) the surface field is no longer connected to the toroidal field (e.g., Schrijver & Title 1999; Schüssler & Rempel 2006), or (2) such connections do exist but the field is highly fragmented and the drag forces are very strong.
the drag to be effective \(v_d < 10 \text{ m/s}\), the transverse size \(a(z_2)\) of the flux elements near the base of the CZ must be a few 100 km or less (see Table 2 in van Ballegooijen 1982). Such a high degree of fragmentation is indeed observed at the solar surface (Stenflo 1989; Solanki 1993; Berger & Title 1996).

The main problem with the disconnection theory is that it is unclear how such a disconnection can be produced. It would require efficient reconnection between the two legs of the \(\Omega\) loop, involving nearly all its magnetic flux (any unreconnected flux would rapidly drift away). Such reconnection may be driven by large-scale flows converging towards the \(\Omega\) loop from the longitudinal direction somewhere in the middle of the CZ. There is presently no observational evidence for such flows. Therefore, in this paper we adopt the view that the surface field remains connected to the deep-seated field but is highly fragmented throughout the CZ. Such a “fibril field” (Parker 1984) is strongly coupled to the turbulent convective flows, causing the field to be passively diffused.

4. Models for the Decay of \(\Omega\) Loops

Flux transport models have been developed to describe the evolution of photospheric magnetic flux over periods of months to years (e.g., Leighton 1964; DeVore, Sheeley, & Boris 1984; Wang, Nash, & Sheeley 1989a,b). These models include the effects of surface diffusion due to supergranular flows, solar differential rotation, and meridional flows. Observations of magnetic helicity in active regions have led to the development of a new class of flux transport models where in addition to the surface flux distribution the full three-dimensional (3-D) structure and evolution of the coronal field is considered (e.g., Amari et al. 1999, 2003; van Ballegooijen, Priest, & Mackay 2000; Mackay & Gaizauskas 2003; Mackay & van Ballegooijen 2001, 2006). According to these models, the magnetic helicity of active regions originates below the photosphere, but the structure of the twisted field is significantly altered by reconnection in the solar atmosphere. The models by Mackay and collaborators use an approximation of the MHD equations in which only the mean magnetic field \(\mathbf{B}_0(r,t)\) is considered and the plasma velocity is assumed to be proportional to the Lorentz force (“magneto-friction”). These models assume that horizontal fields above and parallel to the PIL cannot submerge below the photosphere.

The 3-D flux transport models provide important insights into the build-up of magnetic shear in decaying active regions and the formation of filament channels. Modeling shows that flux cancellation has the effect of concentrating magnetic shear at the PIL. As an active region decays, more and more magnetic shear builds up above the PIL, and the axial flux of this sheared field increases relative to the photospheric flux of the region. Eventually the sheared field can no longer be held down by the overlying coronal arcade and the field erupts, removing some of the magnetic shear from the corona. Generally only a fraction of the axial flux erupts, consistent with observations of “partial” eruptions (Gibson & Fan 2006). After the event the magnetic shear continues to build up, leading to other eruptions at later times. I conclude that 3-D flux transport models provide a natural explanation for eruptive phenomena such as flares and coronal mass ejections (CMEs).
Life and Death of Active Regions

Figure 2. Simulation of the decay of an Ω loop: (a) day 1, (b) day 10. The thick lines are selected magnetic field lines. The thin curves are contours of the photospheric flux distribution. From van Ballegooijen & Mackay (2007).

In order to provide a more complete description of the evolution of Ω loops, it is necessary to develop models for the coupling between the CZ and corona (Amari, Luciani, & Aly 2004, 2005). Recently, van Ballegooijen & Mackay (2007) extended the 3-D flux transport model into the CZ. The computational domain extends from the base of the CZ \( (r = 0.72 R_\odot) \) to some height in the corona \( (r = 2.41 R_\odot) \) where the field becomes radial. Only the decay phase of the Ω loop is considered. A key feature of the model is that it takes into account the sharp transition from a space-filling field in the corona to a fibrous field in the CZ (Parker 1984). The transition layer is assumed to be very thin. Analysis of the magnetic stresses on this layer shows that the mean magnetic field is discontinuous, \( B_{h0,\text{cz}} = f B_{h0,\text{cor}} \), where \( h0 \) refers to the horizontal component of the mean field and \( f \) is the magnetic filling factor in the upper CZ (van Ballegooijen 1997). As a result of this stress balance the mean field in the upper CZ is nearly radial.

In the numerical model only the mean field is simulated. Magnetic diffusion of the mean field is assumed to occur both in the CZ and in the corona, but special types of diffusion are used that conserve magnetic helicity. In the CZ the field is assumed to be passively advected and diffused:

\[
\frac{\partial A_0}{\partial t} = v_0 \times B_0 - \eta_2 (\nabla \times B_0)_\perp,
\]

where \( A_0 \) is the vector potential \( (B_0 = \nabla \times A_0) \), \( v_0 \) is the velocity, and \( \eta_2 \) is magnetic diffusivity in the CZ. Here \( \perp \) is the component perpendicular to the mean magnetic field (Bhattacharjee & Yuan 1995). We use \( \eta_2 = 600 \text{ km}^2 \text{ s}^{-1} \), consistent with the observed dispersal rate of active regions (Sheeley et al. 1983) and the evolution of the Sun’s polar magnetic fields (Wang et al. 1989a,b). The velocity \( v_0 \) includes solar differential rotation and downward convective pumping (Tobias et al. 1998, 2001; Dorch & Nordlund 2001); the assumed pumping velocity \( v_r = -21 \text{ m/s} \). In the coronal part of the model we use a different type
of magnetic diffusion:

$$\frac{\partial \mathbf{A}_0}{\partial t} = \mathbf{v}_0 \times \mathbf{B}_0 + \frac{\mathbf{B}_0}{B_0^2} \nabla \cdot \left( \eta_4 B_0^2 \nabla \alpha_0 \right),$$

(2)

where \(\alpha_0 \equiv (\mathbf{B}_0 \cdot \nabla \times \mathbf{B}_0)/B_0^2\) is the so-called current helicity, and the term with \(\eta_4\) describes hyper-diffusion (Boozer 1986; Strauss 1988); we assume \(\eta_4 = 10^{11}\) km\(^4\) s\(^{-1}\). The plasma velocity is assumed to be proportional to the Lorentz force, \(\mathbf{v}_0 \propto (\nabla \times \mathbf{B}_0) \times \mathbf{B}_0\), so that the coronal field remains close to a force free state (van Ballegooijen et al. 2000). The above equations are solved numerically on a 3D grid in longitude \(\phi\), co-latitude \(\theta\), and radius \(r\) (for details, see van Ballegooijen & Mackay 2007).

Figure 2 shows results for the evolution of a single \(\Omega\) loop. The loop is assumed to emerge from a toroidal flux rope located at latitude \(20^\circ\) just above the base of the CZ. The flux rope has left-helical twist with pitch length \(L \approx 0.5 R_\odot\). The distortion of the toroidal flux rope into an \(\Omega\) shape is done as part of the initial setup and is not explicitly simulated. The \(\Omega\) loop is given a small tilt in agreement with Joy’s law. Figure 2a shows the configuration on day 1 of the simulation when the coronal field has reached a force free state. At this stage, the coronal field is twisted about a quarter turn, consistent with the fraction of the toroidal flux rope that emerged into the corona.

Figure 2b shows the field on day 10 of the simulation. Note that the surface flux has spread out as a result of subsurface diffusion (coronal diffusion has no effect on the photospheric flux distribution). In the upper CZ the magnetic field is nearly radial; this is a consequence of the small filling factor of the magnetic fields in the CZ and the stress balance with the corona. On day 10 about 20% of the surface flux has disappeared from the photosphere and submerged below the PIL (see Fig. 2b). The submerged loops are the result of inter-diffusion of the oppositely directed radial fields in the two legs of the \(\Omega\) loop (downward pumping plays only a minor role in the transport). The loops eventually move down to the base of the CZ, partially restoring the toroidal field.

As found in earlier calculations (e.g. Mackay & van Ballegooijen 2006), the coronal field has become highly sheared. This can be attributed to field-line reconnection that occurs just as the field submerges below the photosphere. The component of the field perpendicularto the PIL is seen to submerge, while the axial component stays behind in the corona. This is due to the fact that the magnetic field in the upper CZ is nearly radial, therefore the submerging field cannot have an axial component. The coronal fields predicted by the present model are similar to those thought to exist prior to CME (Moore et al. 2001; Moore & Sterling 2006). Indeed, the build-up of magnetic shear in our model eventually leads to a loss of equilibrium of the coronal field, resulting in an eruption on day 20 of the simulation. The degree of twist of the pre-eruption coronal field depends on the amount of coronal magnetic diffusion. In models without coronal diffusion, a twisted “coronal flux rope” is formed prior to eruption (e.g., Mackay & van Ballegooijen 2001; Amari et al. 1999, 2003), whereas in models with coronal diffusion such as the one presented here, the pre-eruption field is highly sheared but weakly twisted.

Interactions between \(\Omega\) loops located on different toroidal flux ropes were also considered (van Ballegooijen & Mackay 2007). Figure 3 shows the magnetic
field on days 1 and 5 of the simulation. On day 1 we have two separate bipoles, but on day 5 new loops have formed in the corona between the two active regions: low-lying loops connecting the follower polarity of the low-latitude active region with the leading polarity of the high-latitude region, and higher altitude loops with the opposite connectivity. These new loops are created by coronal magnetic diffusion, and are strongly sheared with respect to the East–West oriented PIL between the two active regions. This suggests that such sites are prime locations for the formation of filament channels (Martens & Zwaan 2001). Magnetic fields submerge not only within each active region, but also in between the two regions. As the flux is transported to the base of the CZ, a new connection is formed between the two toroidal flux ropes. Therefore, as predicted by Martens & Zwaan (2001), the structure of the subsurface fields has been changed by the interactions between the two \( \Omega \) loops.

For the particular models shown above, photospheric flux cancellation is associated with submergence of \( \Omega \) loops (Zwaan 1987, 1996), not the emergence of U-loops (Spruit, Title, & van Ballegooijen 1987; Wilson, MacIntosh, & Snodgrass 1990). However, U-loops can occur in such models when the legs of the \( \Omega \) loop are forced together at larger depth in the CZ. Figure 4 shows results for a model in which the legs of the initial \( \Omega \) loop are initially more nearly radial than in the case of Figure 2, and the diffusivity \( \eta_2 \) increases with depth in the CZ (from 600 km\(^2\) s\(^{-1}\) at the solar surface to 1050 km\(^2\) s\(^{-1}\) near the base). In this case there are some U-shaped field lines below the PIL that are disconnected from the toroidal flux system. Note that the U-loops are highly twisted.

5. Conclusions

The magnetic helicity in active regions originates below the photosphere, but the spatial distribution of the magnetic twist/shear is significantly altered by
reconnection processes in the solar atmosphere. Three-dimensional flux transport models have shown that weakly twisted coronal flux ropes can be formed by reconnection associated with photospheric flux cancellation. The diffusion of magnetic elements on the photosphere transports the footpoints of coronal loops to the PIL, causing a concentration of magnetic shear at the PIL. The sheared field cannot submerge because of conditions in the upper convection zone where the field is nearly radial. Therefore, the photosphere represents a barrier for the submergence of axial fields at the PIL. Magnetic shear builds up in the corona above the PIL, until a certain threshold is reached and the field can no longer be held down by the overlying arcade. Then the field loses equilibrium and erupts, ejecting magnetic helicity into the heliosphere. I conclude that 3-D flux transport models provide a natural explanation for the occurrence of flares and CMEs.

The rate of spreading of surface flux in the 3-D flux transport model with a CZ is consistent with the rate predicted from an equivalent surface diffusion model. This is not surprising because the mean field in the upper CZ is nearly radial. I conclude that surface diffusion models (e.g., Leighton 1964; Wang et al. 1989a,b) are quite adequate when only the surface flux distributions are required.
Furthermore, the evolution of the coronal field in the model with a CZ is very similar to that found in the equivalent model without a CZ (i.e., with surface diffusion and differential rotation applied at the solar surface, and hyper-diffusion in the corona). Therefore, after an $\Omega$-loop has fully emerged into the corona, the further evolution and eruption of the coronal field is nearly independent of subsurface conditions.

In the models presented in Figures 2 and 3 the subsurface fields remain connected to the toroidal flux system. At present, there is no direct observational evidence for such connections. In Figure 4 some of the subsurface field is disconnected from the toroidal field (U-loops), but most of the subsurface field is still connected. It is difficult to see how all of the surface flux could become disconnected, as suggested by Schrijver & Title (1999).

The 3-D transport models predict that submerging magnetic fields are transported back to the base of the CZ, where they “repair” some of the damage to the toroidal field from $\Omega$ loop formation. This is important because the Sun only has a limited amount of toroidal flux, and without such repair too much toroidal flux would be removed by $\Omega$-loops (van Ballegooijen 1997). We found that new magnetic connections can be formed between neighboring toroidal flux ropes. The North–South component of this field is opposite to that which produced the toroidal flux ropes. Therefore, such connections play a key role in the reversal of the poloidal field of the solar dynamo (Zwaan 1996; Martens & Zwaan 2001).

References

Acton, L. W., et al. 1992, Sci, 258, 618