

PERCOLATION AND THE SOLAR DYNAMO

KENNETH H. SCHATTEN

a.i. solutions, Suite 215, 10001 Derekwood Lane, Lanham, MD 20706

Received 2006 June 30; accepted 2006 October 18

ABSTRACT

The origin of magnetic field sources in the Sun’s dynamo is central to this paper. The Babcock-Leighton dynamo was originally envisaged as a shallow dynamo. The source of the Sun’s magnetism is now generally thought to reside near the base of the convection zone and that these fields rise by buoyancy to initiate sunspots. We reconsider this aspect of the solar dynamo. We do this by considering two surface effects as an alternative to the deep origin of the Sun’s magnetism. They are (1) small-scale convective overturning forming the magnetic carpet of ephemeral regions, and (2) percolation, a process wherein the small structures combine to form larger entities. We discuss these effects, and we develop a numerical percolation model and a set of simplified Leighton-type dynamo equations. The numerical percolation model, initiated with two separate random distributions of unipolar fields, does simulate fields clumping together into larger sunspot-like structures, but does not yet display the bipolar nature of actual sunspot structures. We provide a set of simplified global dynamo equations illustrating the temporal behavior of the current percolation model. With the current model being predominantly illustrative, it is envisaged that more realistic shallow solar dynamo models will be forthcoming. We end by providing three types of observations that may distinguish the percolation model from the deep-seated field origin dynamo models. They are (1) the temporal development of activity centers, (2) the magnetic flux distribution within groups, and (3) velocity flow patterns, near and within active regions. In addition, our modeling suggests that a long-term accounting of the amount of flux in ephemeral regions may lead to long-timescale predictions of solar activity.

Subject headings: MHD — solar-terrestrial relations — Sun: activity — Sun: general —
Sun: magnetic fields — sunspots

1. INTRODUCTION

As a prelude to discussing the solar dynamo, let us consider Aristotle’s reasoning (see Webster 2004), as it applies to the source of rivers: “It is clear then that we must not suppose rivers to originate from definite reservoirs: for the whole earth, we might almost say . . . that rivers rise at the foot of mountains proves that a place transmits the water it contains by gradual percolation of many drops, little by little, and that this is how the sources of rivers originate.” Thus, the origin of flows may, at times, be aided by the gradual accumulation of substances, as opposed to originating from some great reservoir. This behavior is also found in the Earth’s atmosphere, where clouds form into their large structure from vapors that have coalesced into aerosols and then into drops. Further along this scale of increasing size and intensity, clouds can then form thunderstorms, and subsequently hurricanes, if conditions are right. In a similar manner, we consider the magnetism in the Sun’s dynamo to originate from small magnetic sources in the photosphere, as opposed to a reservoir in the Sun’s interior. It does not appear that pure reasoning, as was fashionable in Aristotle’s day, can really answer any of these questions. We now recognize that observations are essential to understanding any physical phenomenon.

Solar physics theory has been heavily dependent on and guided by observations of our closest star. Studies have ranged from the smallest observable scales to the global Sun, and in recent decades we have seen into the interior, through helioseismological inferences. Although many advances have been made in our understanding of the solar dynamo, there is still no completely satisfactory model. This gap is made evident from the many disparate estimates for the size of the next solar cycle predicted (e.g., see Svalgaard et al. 2005; Schatten 2005; Dikpati et al. 2006).

Despite this gap in a scientific consensus about solar activity prediction, let us mention some general understandings of the Sun’s dynamo where there is consensus. These predominantly were garnered from early observations of sunspot positions and magnetic fields, namely, the Maunder-Spoerer butterfly diagram, Hale’s laws of sunspot polarities, and the tilt of bipolar magnetic regions (BMRs), often called Joy’s law. Babcock and Leighton put together a rather nice paradigm that allowed an understanding of these observed solar activity features, as well as the out-of-phase oscillation in the Sun’s polar magnetic field, and why the solar cycle should be considered a 22 year cycle as opposed to an 11 year cycle.

Although Babcock originally pictured a rather shallow solar dynamo, Leighton also considered the possibility of a more deeply seated origin for solar magnetism, although he modeled both concepts. Their concepts were concerned with two effects. The first is the omega effect (using the Sun’s differential rotation to create toroidal field from poloidal field). The second was the alpha effect; it arose when one considered how the poloidal field was created. Babcock imagined a shallow field, but the view of a deeply seated source for sunspots was envisioned by Parker and has become prevalent. In this view, magnetic fields are dragged into the interior, most commonly to the base of the convection zone, or the tachocline, and as a result of magnetic buoyancy this deeply seated field can later rise to the Sun’s surface to form sunspots, when the field increases in strength due to stretching. Parker (2001) and others have noted a number of problems in this picture, which have not been overcome. One interesting aspect is how various models explain Joy’s law. The Babcock-Leighton picture provided an explanation in terms of shallow fields. With deep fields, however, a separate mechanism seems to be required to explain this. Nandy & Choudhuri (2002) and Longcope & Choudhuri (2002) have

developed models to explain the tilt of active regions. This tilting seems to be an interesting difference between the deep and shallow models of the Sun's dynamo.

The changing paradigm away from the shallow dynamo of Babcock, to the modeling of both shallow and deep dynamos that Leighton explored, and now to the more modern picture of predominantly examining deep solar dynamos, occurred for the following reason. It was thought that shallow dynamos have a theoretical problem. Namely, a horizontal line of magnetic field is thought to be unstable; it would rise quickly and dissipate owing to its magnetic buoyancy. As a consequence, it became fashionable to simply place the origin of the Sun's dynamo magnetic field deeper, somewhere near the base of the convection zone, close to the stable layers of the radiative core. Ruzmaikin (2001) discusses many observational and theoretical aspects of the deep dynamo, including clustering, closely related to the percolation aspects we discuss. This deep dynamo allowed a possible alternative to the shallow magnetic buoyancy problem. A shallow dynamo may still be possible; let us discuss this.

The problem associated with the persistence of shallow field lines may be overcome by the combined effects of the following properties associated with solar magnetism in the Sun's convection zone. As the footpoints of a newly generated shallow flux tube are separated, field stretching and increased tension encourage the magnetic tube to sink (as a chord does on a sphere), relative to the buoyancy associated with a static hydrodynamic force. Naturally, there is interplay, since hydrodynamic forces provide the energy for magnetic field growth. In any case, a field line would not likely be totally level over an extended length, but only at a finite number of points; thus, one must consider the forces on tilted submerged field lines. If one considers a finite element of such a field line, driven by a downflow, then this is equivalent to the problem of the orientation of a general object in a fluid flow. The solution to this is well known: rather than aligning with the flow, the object is reoriented so that it is perpendicular to the flow force in such a manner that it subsides with its bowed surface oriented perpendicular to the flow (e.g., the bowed shape of a leaf points "downward" as it falls from a tree, essentially floating, thereby maximizing, rather than minimizing its drag). Thus, a field line descending from an ephemeral or active region, rather than pointing radially inward, has a tendency to be deflected away from the radial direction and is reoriented toward the horizontal rather than extending toward deeper layers. This occurs until it can extend outward toward another footpoint. Last, a flux tube is not an isolated field line in the presence of a static fluid; instead there are turbulent forces, not simply static pressures, and the background does not contain zero magnetic field, as nonzero magnetism is ubiquitous. Thus, the motion and stability of flux tubes is governed by the complex forces of the magnetic tension associated with the field's footpoint motion, along with the turbulent pressures from the complex surrounding media. This is not a closed problem.

In addition to the overall global picture of solar magnetism, individual sunspot physics is also an intriguing subject. The manner in which sunspots are held together by fluid dynamical forces was examined by Weiss, wherein the inflow of fluid provided a mechanical "collar," which held the magnetic field together. Energy transport was also important—how do sunspots remain cold against the onslaught of convective energy transport? The Biermann picture of the magnetic field inhibiting convective energy transport was augmented by Parker (1979), who used the Weiss flow pattern, and showed that a "superadiabatic effect" could aid sunspots to remain cool, against the tremendous radial outflow of convective energy present in the upper levels of the Sun's convection zone.

A simple picture may aid understanding, if we consider the normal outflow of energy as a "river of energy," due to the high flow rate, compared with the reservoirs of energy storage available. Sunspot fields can only prevent the outflow of energy for only modest time periods, in the layers of the upper convection zone, as there is little ability to store significant energy in these thin layers. Examinations of the outflow of energy from the Sun come from the work of Willson et al. (1981), Lean et al. (1982), Chapman et al. (1986), and Sofia et al. (1982). It has been shown that sunspots are associated with a solar constant "energy deficit" and bright faculae provide an offsetting excess. More recently, we have seen that the contribution of bright faculae actually supersedes the sunspot deficit, providing an excess of energy released in the Earth's direction associated with active regions. Thus, the effect of solar activity is actually opposite to the effect of sunspots, with an activity center releasing somewhat more energy, rather than less (Lean et al. 1982). These modern advances in solar activity deal with the effect of solar magnetism on energy transport in the Sun's outer convection zone. The excess energy emitted is important for this paper, as it helps explain why sunspots are drawn together, as explained later.

In addition to how sunspots affect the solar irradiance, there has been a long history of observations of sunspot development dealing with their birth, growth, and decay. Despite this, our knowledge of sunspot evolution stems primarily from photospheric observations and thus their full three-dimensional structure is uncertain. It is because of this, that we feel obliged to raise and consider an alternative picture to the modern, deep-seated origins of sunspots. The relationship of solar magnetism to sunspots is difficult to fully understand, owing to the complicated field geometries that may arise. Until recent helioseismological interpretations, we have only been able to view the three-dimensional field in one surface, the photosphere, and often only the line of sight component; hence, the three-dimensional field geometry and structure have remained elusively hidden within the Sun. Although helioseismology has allowed the photospheric cloak to be unveiled, we still have not been able to trace field lines well into the Sun's interior. Thus any paradigm for sunspots must undertake a double duty: it should correctly fit into the magnetic cycle of the Sun's dynamo and also each sunspot's individual life cycle should be adequately portrayed.

Traditionally, it has been noticed that bright faculae move away from sunspots in their late, decay phase. Harvey & Harvey (1973) focused on these moving magnetic features (MMF) that surround sunspots like a moat. Their three-dimensional structure was still uncertain; however, they showed up as "bright points" in the high-resolution observations, making them faculae in the traditional sense. In addition, they seemed to carry off magnetic flux from sunspots, and thus were correlated with the decay phase of sunspots. Indeed the Harveys focused on the late phases of sunspots' lives. Hagenaar & Shine (2005) find the same general behavior with greatly improved observations. We discuss below how field lines can be expelled from active regions, consistent with these observations.

The picture of a fairly deep origin to the Sun's dynamo has been opposed by Brandenburg (2005), who advocated a shallow solar dynamo. Babcock (1961) stated in the first line of his abstract "Shallow submerged lines of force . . . produce a spiral wrapping of five turns after . . . three years." Leighton developed both the deep and shallow dynamo models, with the deep model providing a better fit, as it allowed more differential rotation to magnify field with. Babcock thought the shallow picture was supported by some of the observed aspects of sunspots. We shall attempt to further this picture of the shallow dynamo, by considering a

new process, percolation, for magnetic flux magnification near the Sun's surface. Schrijver (2001) has also provided simulations of the Sun's dynamo via surface features, with ephemeral regions, differential rotation and meridional flow patterns. These are more detailed than our dynamo model as we try to illustrate the importance of combining shallow fields (via percolation) and thus have left out many elements that Schrijver and others use to describe the real solar dynamo.

In a nutshell, the basic idea we put forth is the following. Magnetic fields originate in the small magnetic carpet bipolar/ephemeral regions described by Schrijver & Title (2001). These are then drawn into newly developing active regions and the energy associated with fluid dynamics couples with the magnetism of these regions to allow a larger, more efficient flow pattern to emerge, as magnetic flux and the structural size increase. Magnification of magnetic fluxes occurs in this view, not by the usual mechanism of stretching magnetic field and thereby increasing its field strength, but rather by the inflow of like-sign magnetic fluxes into the developing active region. This increases the magnetic energy for it takes energy to force like magnetic fields together. This is obvious to anyone who has played with magnets and also from Maxwell's equations. This energy increase, in the current view, is powered by the large-scale flow patterns via cool downflows below strong fields that pull together same-sign field regions, through lowering the total energy (field, flow, heat, etc.). The cool downdrafts allow a larger scale, more efficient flow pattern to exchange energy outward from the solar interior (see Schatten & Mayr 1985). The excess emitted energy, as discussed before, plays an important role in the process. This occurs only in areas of the Sun where the superadiabatic gradient is large. We presume that near the surface of the Sun, where the gases are thinner, they are subject to more fluctuations in energy/temperature than in the deep interior. With this in mind, we consider that in the shallow layers below the Sun's photosphere there are areas that are hotter and cooler than the average structural value. Areas where the superadiabatic gradient is largest, in our view, would be particularly conducive to sunspot growth. After a few days, surrounding areas allow cool inflows to form leading to cool downdrafts below the sunspot and can quench the superadiabatic gradient thereby allowing the sunspot to decay. One may refer to the growth phase of this process as the "superadiabatic magnetic flux magnification" process, but here we substitute the simpler but less descriptive term, percolation. We will discuss the process more fully, but it is extended from Parker's (1984) model explaining why like magnetic elements in convective cells are drawn together.

Percolation is essentially the term given to the physics and mathematics dealing with the spatial connections of features. Examples of percolation deal with clustering, criticality, diffusion, fractals, phase transitions, and disordered systems (see Stauffer 1994). Thus it is used to describe the processing of small features into a larger structure. Most people are more familiar with percolation from the name of older coffee makers combining bubbles that push boiling water through a filter containing coffee grounds. Percolation is a noun of action, derived from *percolare* ("to strain through, filter"). For the solar case, we use percolation to refer to the processing of small ephemeral "like-sign field regions" into larger features. In our model, percolation in the Sun is thought to allow small-scale fields (ephemeral regions or pores) to combine into larger sunspots in the upper convection zone, through the extra available energy in the Sun's upper convection zone, which is highly superadiabatic.

For our purposes, we are dealing with percolation of magnetic fluxes in the solar photosphere. We shall be invoking the percolation process such that only fluxes of the same sign are prefer-

entially drawn together, due to a percolation force we discuss. We later consider how the dynamo equations of Leighton might be modified to incorporate our present view, and then some observational tests, that could be conducted. The current viewpoint basically suggests an important role of the near surface region (from the photospheric boundary to the hydrogen ionization region, at a depth of roughly 20 Mm) for the Sun's dynamo.

2. EPHEMERAL REGIONS AND PERCOLATION

We shall be discussing two processes that may support a shallow solar dynamo. The first refers to the germination of magnetic field within ephemeral regions (EPRs) and the second; we refer to as percolation, the workings of which we explain below. The first process has been well observed recently with powerful telescopes from the ground and from space, and is understood theoretically (e.g., see Schrijver & Title 2001). Observationally, these small regions are associated with large magnetic field and high-energy (effective temperature) plasmas, as seen in the EUV, soft X-rays, and other high-temperature electromagnetic emissions associated with the Sun's upper atmosphere. We simply describe the main behavior and refer the reader to numerous articles on the subject. It occurs as convection (e.g., granulation) wraps and separates small-scale fields into tight structures. This binds and isolates the features into two diverse structures: (1) small regions of high field strength and low turbulent velocities, and (2) larger regions of much lower field strength and higher turbulent velocities. This germinates ephemeral regions: small-scale bipolar fields formed from convective overturning of the Sun's magnetized plasma. The process is ubiquitous all over the Sun, both in location and time, throughout the solar cycle. The process has been studied by many in the Lockheed group and others. In their view, the "magnetic carpet" name was chosen for the myriad of small loops that these features form above the Sun's surface. Thus, they are important in the structure and heating of the Sun's corona. Parnell (2002) discusses some of the numerical aspects of how much flux is generated.

The second process is more novel; we refer to it as percolation, as physicists have used this name for the processing of small features into larger structures, primarily in solid state physics. In our model, percolation near the photosphere is thought to allow small-scale fields (EPRs/pores) to combine into larger sunspots through the available energy in the Sun's upper convection zone, which is highly superadiabatic. Following Parker's (1984) work, where convective elements isolate magnetic field by lowering the convective flow energy, his process suggests to us, that in a superadiabatic ionized atmosphere, magnetic fields of like sign attract, and unlike fields repel, essentially the opposite behavior of magnetic fields in a vacuum, or subadiabatic atmosphere. Thus our model invokes a stickiness to the same sign (same sign radial component) magnetic fields in the photosphere. This stickiness or percolation thrives from lowering the free energy in the presence of convective or superadiabatic energy transport. Thus, the percolation process allows the Sun to shed its luminosity more efficiently (than a totally unmagnetized, convecting atmosphere) by having flow transport energy in uni-directed (up-down) flows rather than convective bubbles that "break" in small distances, comparable to the atmospheric scale-height, as in mixing length theory. Let us now examine the percolation process in more detail, so as to better understand the origins of this.

To understand the difference of flux tube behavior within the Sun's convection zone, and conventional magnetic field behavior, we can do no better than to quote Parker (1984), who states: "Now in ordinary turbulence, the field would grow (at the expense of kinetic energy) only until the Maxwell stresses became

comparable to the Reynolds stresses, whereupon the two systems strike up a dynamical balance of some sort . . . But the convectively driven turbulence in the Sun represents a different situation. The turbulent mixing cannot be suppressed by the magnetic field because the mixing is so necessary to transport the energy from the radiative core to the surface of the star. The convection cannot slow, or become oscillatory, merely because it is hampered by the growing magnetic stresses. . . . The question is what happens? If the convection continues to mix fluid, to transport heat, then the field lines continue to be stretched and the field intensified, requiring an increasing temperature gradient (caused by an accumulation of heat below) to force the turbulence mixing against the growing Maxwell stresses. Precisely how severe this confrontation between Maxwell stresses and convective forces might become is difficult to estimate. . . . Assuming that the gas beneath the convective zone is largely free of field, the pile up of heat forces columns of extra hot field-free gas upward through the tangle of magnetic field and stalled convection. The magnetized gas is pushed aside by the overheated field-free gas from below." Parker then shows that the field would increase to a peak fibril field value such that $B_f^2/8\pi$ is a fraction of the gas pressure $p(0)$, independent of the mean initial field value $\langle B \rangle$. Parker then calculates the values of magnetic fluxes for vertical and horizontal fibrils. The important point is that the field strength is forced to rise from a mean field value $\langle B \rangle$, to a fibril field value independent of this, B_f so that the field can be sequestered into a small volume of space in such a manner that it no longer inhibits the convective energy transport. It does this to ensure that the turbulent convection can still transport the same energy flux. Parker calculates the value, B_f so as to minimize the energy of the system (the sums of thermal gravitational and magnetic energies). Even though the magnetic energy rises, it is a small price to pay to reduce the other forms of energy. The thermal term is a continuous energy form, which if not allowed to transport energy, would build up continuously. Thus it cannot be defeated by the magnetic flux.

We make implicit use of Parker's theory that fields lower the overall energy when like signs combine into small structures, and are thus sequestered away from the remaining area. Not having a detailed model of the changes in flow energy with field sign, configuration, etc., we simply invoke a force for larger areas than a convective cell, which may be referred to as the "superadiabatic magnetic flux force," or more simply the percolation force, that we assign to magnetic field regions that would cause them to behave in the manner suggested by Parker's model. This percolation force would be *opposite* to the normal vacuum or subadiabatic magnetic force between two poles, i.e., it would cause like poles to attract and unlike poles to repel. This force may be represented by a potential that would cause a confluence of like field in the photosphere. Since convection is controlled by the superadiabatic gradient, S , it would be proportional to S .

We choose the percolation force, F_P , between two field sources, 1 and 2, by modifying the hydrodynamic force, F_H , as follows:

$$F_P \approx F_H \frac{B_1 B_2}{|B_1 B_2|} S \approx \frac{\rho v_0^2}{2} \frac{B_1 B_2}{|B_1 B_2|} S, \quad (2.1)$$

where the hydrodynamic force, F_H , is approximated by using the expression for the force on a Pitot tube embedded in a horizontal flow of velocity, v_0 . The "forces" in equation (2.1) actually represent pressures (forces per unit area). Figures 1a–1d shows this attractive force behavior of like fields when the superadiabatic gradient is positive. Driven by the convective energy

flux, small disturbances in the surface (magnetic fields of the same sign) are attracted allowing these field structures to grow, at the expense of the superadiabatic energy. Figure 1e shows the reverse process. Downdrafts below sunspots (as discussed by Schatten & Mayr 1985) cool the hot material below the sunspot as it ages; this reverses the superadiabatic gradient, and the percolation force then expels same-sign flux from a sunspot. The percolation force, being dependent on the superadiabatic gradient is important in the life history of sunspots, as mentioned in the introduction. After the growth phase of a sunspot, the cool downdraft quenches the heat below the sunspot and can reverse the superadiabatic gradient. This then reverses the percolation force and the heat flow and allows like-sign fields to repel and faculae to form. This in turn results in magnetic fields being expelled from sunspots in their late phases. This is consistent with the observations of Harvey & Harvey (1973) and Hagenaar & Shine (2005).

Parker's work dealt with like-field regions in a granular cell. Let us discuss why we think the force of attraction of like fields, and a repulsive force of unlike fields, extends to larger regions. First, we consider the larger scales. In general the convective energy transport (see Cox & Giuli 1968; Mihalas 1970; Schatten 1988) may be written as

$$F_{\text{conv}} = \rho C_p K S, \quad (2.2)$$

where ρ is the density, C_p the specific heat, K the eddy thermal diffusivity, and $K = \frac{1}{2} v_t l$, where v_t is the turbulent velocity, l is the scale of mixing, and S is the superadiabatic temperature gradient, $S = |dT/dr| - |dT/dr|_{AD}$. When this is positive, convection transports energy outward. Equation (2.2) shows the convective energy transport is proportional to S but also the thermal diffusivity, K . This factor grows with the scale length of the mixing (or turbulent diffusion) length. Thus, with larger scale mixing (than in a small convective cell), the length is larger and with it, the outward energy transport is enhanced. The effect of increasing the convective energy transport is to lower the total energy of the fluid. This is actually seen as the Sun radiates more energy with higher solar activity (see Lean et al. 1982; Willson et al. 1981). Considering this factor as more important than the magnetic energy, as in Parker's model, provides that in a superadiabatic environment, the same sense field is attractive (as is observed in sunspot growth, for example, when the superadiabatic gradient is positive, energy is convected outward, and this free energy allows magnification of small disturbances).

Now we consider why fields of opposite sense may not simply cancel, but in a superadiabatic environment, may separate (repel) each other. Let us consider the following thought experiment, with two field geometries. In addition, we accept that same sense fields are attractive at some locations on the Sun, such as sunspots since their magnetic fields are seemingly held together by the flow field. Now, in our first geometry, consider two same sign field sources, say A and δ , with total energies, E_A and E_δ , respectively, in such an environment. As earlier discussed, when added together one obtains a reduced total energy: $E_{A+\delta} < E_A + E_\delta$. For our second geometry, we consider opposite fields, so the addition of a field of opposite sign (from A), say $-\delta$, of the same magnitude as δ but of opposite sense, is added to a larger field, $A' = A + \delta$. This field would then add as follows (to the larger field of $A' = A + \delta$), thereby returning its energy to its previous larger energy. Basically, one simply adds $E_{-\delta}$ to the previous inequality, and obtains $E_{A+\delta} + E_{-\delta} < E_A + E_\delta + E_{-\delta} = E_A$. Thus, $E_A > E_{A'} + E_{-\delta}$, so two oppositely oriented fields, in this scenario repel (in the solar superadiabatic environment where sunspots grow).

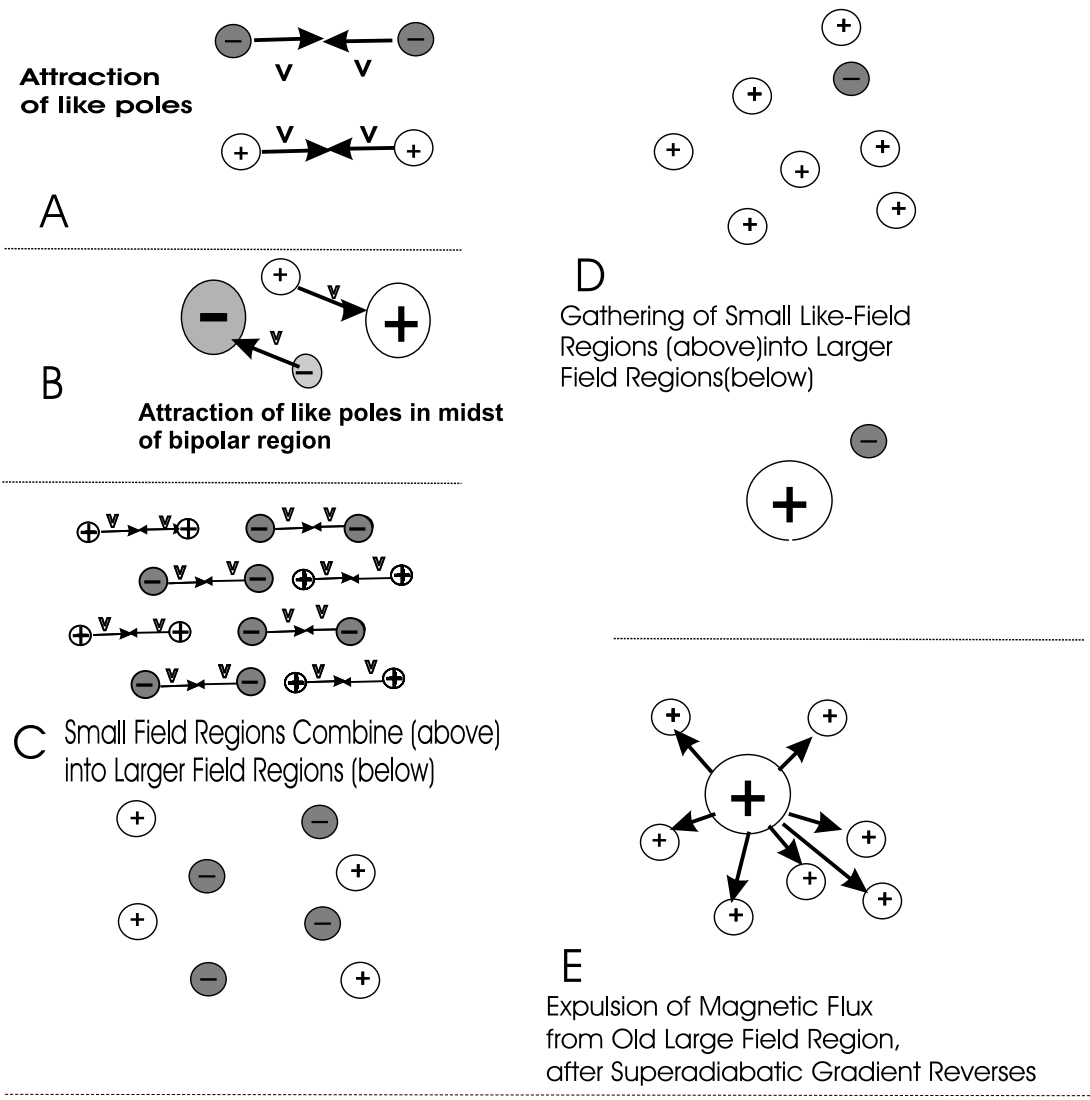


FIG. 1.—(a–e) Top-down views of photospheric magnetic fields. In this figure the superadiabatic gradient, S , is positive. In these cases, like-sign fields attract due to the flow patterns as shown by the arrows, indicating motion. Panel e shows the expulsion of magnetic flux from a sunspot region when the region ages, and the superadiabaticity, S , reverses.

The role of flow on active region energy transport was discussed by Parker (1979) and later by Schatten & Mayr (1985). They suggest sunspot growth occurs when S is large, and a subsequent downflow beneath sunspots can quench the process after the free energy is depleted locally, comparing these features to an ion hurricane. The analogy is formed, based on the ionization energy in the Sun's atmosphere playing a similar role to latent energy of water vaporization in the Earth's atmospheric weather systems. For sunspots the magnetic field organizes the large-scale flow into inflow and downflow patterns, and for the Earth's atmospheric hurricanes, the presence of large low-pressure regions plus Coriolis forces provide organizing effects. For sunspots, a surrounding and subsequent upflow is thought to release the energy offset (from the neutral hydrogen downflow) into faculae, which are bright features surrounding sunspot groups. Overall, the effect of sunspots and faculae on the solar constant is positive, meaning more energy is released as a result of the active region development over the lifetime of these regions. Thus, the net effect of active regions helps the Sun to shed its luminosity, and this is consistent with the increased efficiency of convection associated with the sunspot increasing the scale (active region size) of energy transport compared with the mixing length (gran-

ular size). Let us illustrate the effect of the percolation force on magnetic features (e.g., pores/EPR), in superadiabatic regions near the photosphere.

The percolation force results in flows as shown by the arrows in Figure 1. We illustrate these flows in the following cases: the attraction of like poles, in a weak field area, and within a bipolar region; the development of small field regions into larger field regions (particularly in regions where the superadiabatic gradient is large; and how small field regions develop within a region of one dominant polarity. These patterns illustrate growth cases, where and when S is large. For small or negative S , fields would decay. Figure 2 displays on the left how combining two separate field structures into one makes the flow pattern larger and more efficient. Figure 2 (right) shows how field and flow patterns may become separated, owing to the magnetic tension and gravitational flow force being directed differently. The cool downflow is directed radially inward, whereas the field tension may point in any direction. In our case, with field thought to exist predominantly in shallow regions, it would turn tangentially in short order. With increasing depth and pressure, the field lines shrink rapidly with depth, and thus may separate from the flow pattern, as the hydrodynamic force decreases rapidly as the area of a flux tube

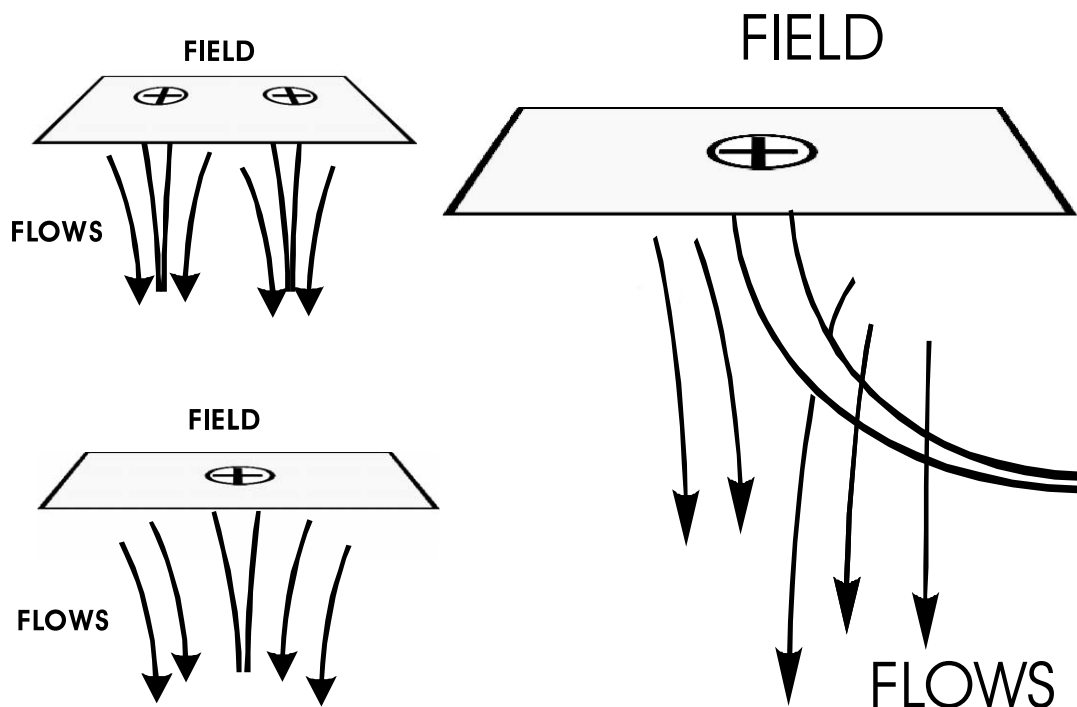


FIG. 2.—*Left:* Combining two separate flux tubes and surrounding flows (*top*) into one larger flux tube (*bottom*), in accord with Parker's model, allows a reduction in the total energy. This increase in efficiency with larger flow size is the origin of the percolation force. *Right:* Although many flow fields are considered symmetrically attached to field fibrils, at a few scale heights deep, the field structure breaks up into individual fibrils (not shown), and the field tension may draw it off to one side allowing it to connect to some other surface field of opposite sign. The cool downdraft may continue to flow inward, driven by gravitational forces, or the flow pattern may break up. The differing behavior of flow and field may allow a separation of the magnetic field vector from the flow vector (pointing in different directions).

is reduced with pressure. In the Figure 3 pictorial, we see the dissociation of flowlines with magnetic field lines, where downflows occur below field regions (pores and spots are usually dark, but for the purposes of this illustration are light), with upflows surrounding them. Magnetic fields below the Sun would connect features of opposite polarity, but hydrodynamic forces could separate or connect field lines from different sources.

Numerical simulations of flux expulsion in a compressible medium were undertaken by Hurlbert & Toomre (1988). They considered the dynamical problem in two-dimensional Cartesian geometry and simulated thermal convection in a closed box, spanning a few pressure scale heights. Compressibility breaks the symmetry between up- and downflows and leads to concentrated, rapid downflows and gentler, extended upflows. As a consequence, most of the magnetic flux becomes concentrated in the downflows. The growing vertical magnetic field suppresses the horizontal fluid motion but does not interfere with the vertical velocity component; therefore, the upper part of a forming flux sheet becomes partially evacuated and then compressed by the external fluid pressure, agreeing with Parker's theoretical arguments. It is hoped that observations may settle whether magnetic fields combine from percolation, as we describe, or arise from deep flux tubes as in the conventional picture. We are currently unaware of any observations that can speak directly to the three-dimensional time history of solar magnetism, however, it is hoped that the proper interpretation of high-resolution, high-cadence observations of growing active regions from *SOHO* MDI or other modern ground-based solar telescopes may be used to address this question.

Having discussed the physical origins of the percolation force, let us examine its influence on photospheric fields. We provide two numerical models to illustrate the effects of the percolation process on solar fields and discuss how this can lead to a surface

origin for the solar dynamo. We first describe the numerical basis for two-dimensional percolation modeling, and then how this might affect the Leighton dynamo equations, using a simplified form that better demonstrates the principles, and the temporal behavior. Following this, we describe some of the geometric behavior of this surface dynamo.

3. PERCOLATION MODELING— NUMERICAL SIMULATION

In the previous section, we discussed ephemeral regions and introduced the basis for percolation as applied to the Sun. Here we describe a percolation model and see its workings in a numerical simulation. One study where percolation was considered for sunspots is the model of Seiden & Wentzel (1996), which had fluxes released from the solar interior simulating sunspots. The paper showed power-law behavior at small sizes and exponential behavior at larger sizes, comparing well with the observations of Harvey (1993). In the present paper, we are ignoring some of these interesting geometric details and focusing only on the broad aspects of how an initial set of small bipolar fields (like ephemeral regions) may combine into larger structures of the same sign and thus allow magnetic flux to separate. This separation has implications for the large-scale field and may give rise to dynamo type behavior, as seen in the photosphere. Thus, we are trying to gain knowledge of the large-scale structures and are examining the question, can larger scale magnetic fields be formed from the accumulation of small-scale features, through percolation? Hence we are attempting to paint a broad background, rather than detailed aspects. Nevertheless, we admire the results of Seiden and Wentzel and are pleasantly surprised at the rich results of their study.

Percolation is essentially the term given to the physics and mathematics dealing with the spatial connections of features. Examples

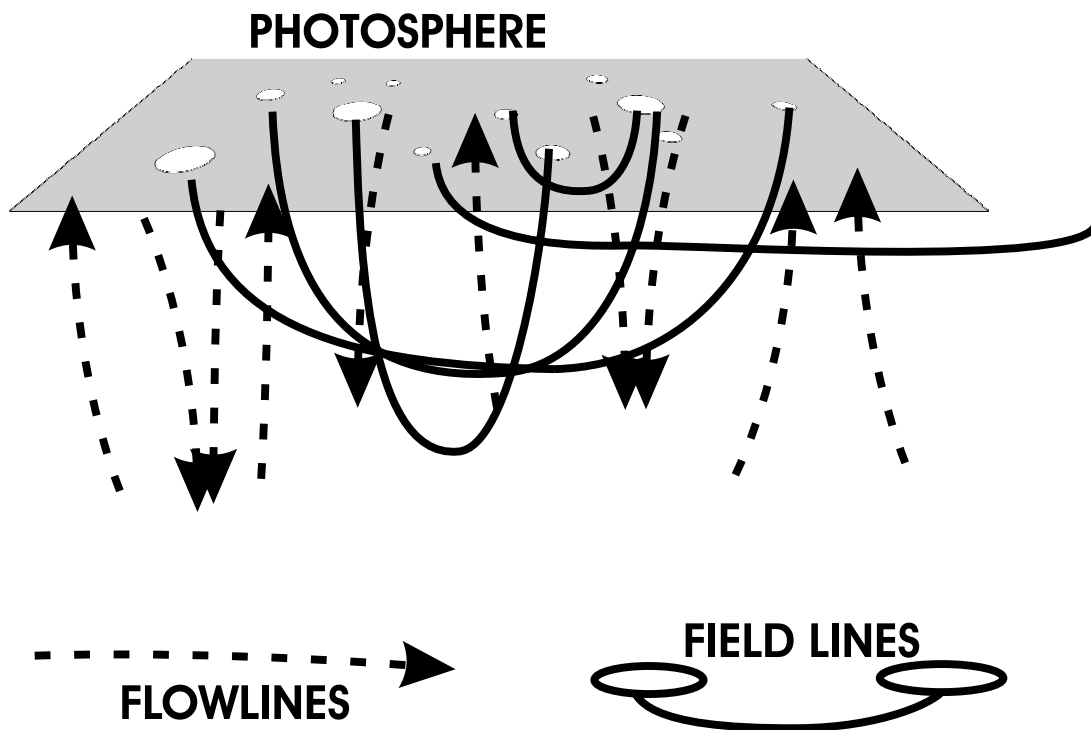


FIG. 3.—As shown on the bottom, flowlines are displayed with dashed arrows, and magnetic flux areas, as circles in the photosphere, connected by solid lines representing the magnetic field without direction displayed. As discussed in Fig. 2, a disassociation between flow lines and magnetic field lines may occur with depth, allowing widely separated fields to connect with one another close to the photosphere. The photospheric field regions are usually dark, such as pores and sunspots, but for clarity in the illustration, they are drawn light. Downflows are thought to occur below sunspots and return upflows surrounding them. Magnetic fields below the Sun would connect features of opposite polarity.

of percolation deal with clustering, criticality, diffusion, fractals, phase transitions, and disordered systems (see Stauffer 1994). Thus, it is used to describe the processing of small features into a larger structure. Most people are more familiar with percolation from the name of older coffee makers combining bubbles that push boiling water through a filter containing coffee grounds. Percolation is a noun of action, derived from percolare, “to strain through, filter.” For the solar case, we use percolation to refer to the processing of small ephemeral “like-sign field regions” into larger features. In our model, percolation in the Sun is thought to allow small-scale fields (EPRs/pores) to combine into larger sunspots in the upper convection zone, through the extra available energy in the Sun’s upper convection zone, which is highly superadiabatic.

This percolation model is calculated using cellular automata, based on the following rules. These rules provide a potential energy term, e.g., as in the energy terms that bring and hold together the magnetic field within pores and sunspots (despite their magnetic pressure tending to separate them), as outlined by Parker, in the presence of the Sun’s superadiabatic region. We use the methods of cellular automata (see Adami 1999, chap. 2.4), and simulated annealing to provide for cell transference, subject to a “condition” or constraint of energy minimization (for simulated annealing [Press et al. 1986, p. 328], a common energy term, is the shortest path distance; for Meyer [2000], it minimizes the total spin energy). This indicates that rather than having the magnetic field move in accord with flows and forces, which could then govern the motion, we simply move surface elements at random by small amounts and have like sign fields “stick.” This is a fairly straightforward but conjectural method that we hope is realistic, however, this aspect remains to be seen. Other algorithms, which share this general basis, are the Metropolis and Glauber methods.

The original initial position has random fields (with three values: 0 and $\pm K$, a constant, similar to Parker’s fibril field value, B_f), such that the nonzero field occupies a fixed fraction (chosen at 10%). For the quiet Sun, the filling factor is thought to be more like 1% (see Lin & Rimmele 1999). The low percentage (10%) is a compromise, chosen to approximate the small filling factor of fields on the Sun, with the desire to illustrate the behavior of the clustering of fields in a figure without vast gaps. The clustering process is also enhanced, with a higher percentage of fields, much as car crashes occur on more crowded roads. Thus, the results are only illustrative, as opposed to representative, of the actual solar surface. The Sun is excellent at sequestering its very large magnetic field to a small fraction of the solar surface, owing to the extremely large surface irradiance it possesses that can tame this field.

Thus, the initial surface field in our model is assumed to emanate by convection, as in the magnetic carpet (see Schrijver & Title 2001) as a result of magnetic pumping, essentially the process that germinates ephemeral regions from convection. To emulate this, we simply start with a null field, and then fill in a small percentage of random signed fields for the system. To achieve field “stickiness,” the percolation properties we desire for field entities in the photosphere, we use cellular automata with the following properties. The cells exist in a rectangular grid, with toroidal geometry to remove edge effects, by having the left and right edges as well as the top and bottom edges connected. Cells have three states, inward ($-$), outward ($+$), and neutral (0). These states are randomly initiated for each cell and do not change. Instead, the cells are allowed to move during each “time interval,” until the simulation ends. For each time interval, a number of cell movements may occur through positional cell swapping. When a cell is considered for swapping, the number of same states, in the immediate neighborhood is calculated. If a cell has

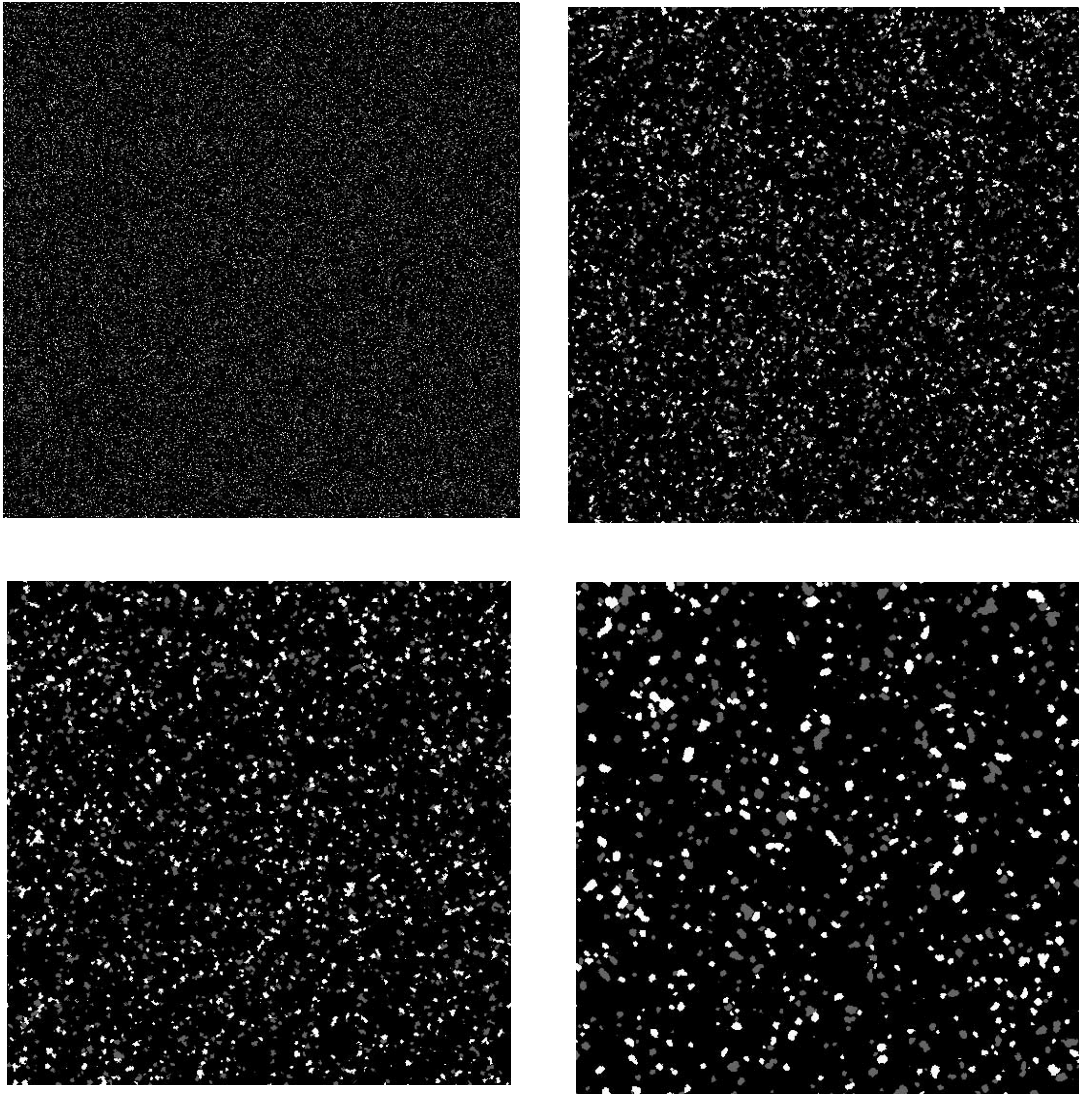


FIG. 4.—Top left displays the initial state with 10% fill, of random fields with two signs, and 90% zero values. The zero values are black, the bright values represent positive values, and gray areas, negative values. Top right displays the result after 3 steps; bottom left, after 7 steps, and bottom right, after 30 steps. Notice that the same percentage (10%) is filled in each figure; however, the field regions are becoming clustered, due to percolation.

all the same sign neighbors, then the cell is internally situated within a pore or spot, and its position is not changed. Otherwise a cell swap may occur based on probabilities for improved placement. The cells are thus able to generally increase the number of neighbors of same sign cells, resulting in the emergence of clusters of cells in the same state. This essentially provides for an effective clustering, or percolation, as in the superadiabatic magnetic force model, previously discussed. Future changes in this methodology may include initiation of field with small dipoles, large-scale field interactions with the cells, fluid motions such as differential rotation, and possibly more complex models such as enhanced flows associated with the field geometry and evolution (growth of regions based on energy and field considerations, etc.).

The first simulation is shown in Figure 4. A set of $\sim 250,000$ cells is chosen in an initial state of 25,000 random bits (of $+/-1$). This is a 10% filling factor. The zero values are black, the bright values represent $+$ values, and the gray areas are for negative values. Top right displays the result after 3 steps; bottom left, after 7 steps, and bottom right, after 30 steps. Notice that the same percentage ($\sim 10\%$) is filled in each figure; however, the field regions are becoming clustered, due to the clustering force.

To illustrate the effect of further steps and also examine things on a smaller scale, we see the illustrations in Figure 5. A smaller scale size, with longer computational steps, allows the computer calculations to be performed in reasonable times. Again, top left is the initial state, with ~ 8000 nonzero values. We now choose jumps of 25, 200, and 1000 steps; bottom left, after 200 steps, and bottom right, after 1000 steps. After 1000 steps, the entire field is in an almost entirely clustered state. Very little happens near the end; the computations only removes two regions, in the last 800 steps. For comparison, in the beginning, there are ~ 8000 groups in this simulation, and 25,000 in Figure 4.

One other point deserving comment illustrates a lack of agreement between these simulations and observed solar behavior. That is the following. The Sun's magnetic fields have not only a particular individual distribution (small regions of high field strength and large regions of near zero strength), but also a large degree of spatial correlation, or spatial coherence. The former relates to filling factor, and this we purposely overestimated for illustrative purposes. The latter relates to how the fields are distributed into active regions, which generally have fields of both signs (bipolar magnetic regions), thus the presence of a large

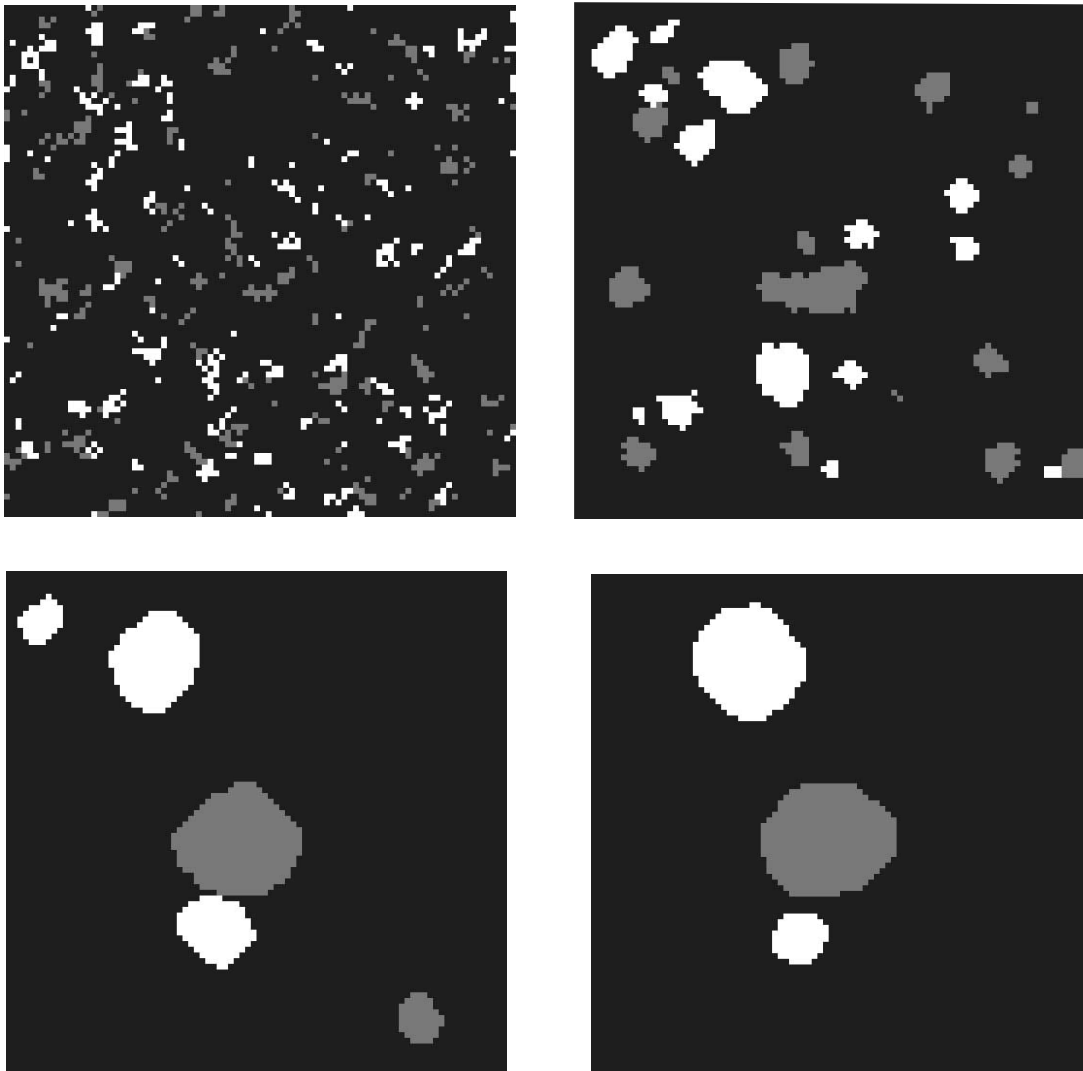


FIG. 5.—Shown is a magnified version of clustering, or percolation. To illustrate the effect of further steps, it helps to show a magnified portion of this simulation. Again, top left is the initial state. Top right displays the result after 25 steps; bottom left, after 200 steps, and bottom right, after 1000 steps. Note that the same percentage of filled field exists, however, at the end the fields are in an almost entirely clustered state, owing to the percolation process.

positive (sunspot) field is often correlated spatially with negative fields (opposite sign sunspots) nearby. Our simulations do not show this. We suppose that our simulations are missing something, but our mechanism (of field percolation) could also be incorrect, and the upwelling of flux loops adequately explains this well-known fact of solar behavior. Our percolation model is highly simplified; it starts with random unipolar fields. Thus, there are two independent distributions of unipolar fields that evolve independently, except they cannot both occupy the same space, rather than starting as close bipolar pairs as ephemeral regions do, owing to Maxwell’s divergence B equation. Thus, our model is simplified, compared with the Sun, by not having highly correlated small bipoles (EPRs), and other aspects that treat percolation in a simplified fashion, rather than how real flows on the Sun could draw actual field structures together.

The Seiden & Wentzel (1996) percolation model does show bipolar behavior, but these authors start with correlated, deep field sources. Another possibility to explain the spatial correlation of active regions has been put forth by Svalgaard & Wilcox (1976). They showed that active regions form straddling sector boundaries in that particular hemisphere where the sector polarity change agreed with Hale’s law for that hemisphere. Their paper suggests

that active regions form at specific places on the Sun’s surface. It is difficult to know how many aspects of the spatial coherence of active regions (e.g., active longitudes) a particular model should achieve. Nevertheless, we feel that the general bipolar nature of active regions is a goal we should strive for in the next improvement of this model. Let us now see how this process might modify the Leighton dynamo equations, by allowing the fields to be initiated through surface effects, rather than arising from upwelling within the solar interior.

4. SIMPLIFIED DYNAMO EQUATIONS

In this section we discuss Leighton-type dynamo equations, which we simplify. The purpose of the modifications is to change some of the field variables (Leighton had radial field and “un-erupted flux”) into variables that are more suited to our view of ephemeral region (EPR) fields and active region fields, using sunspot number, R_z , as a representative index, although any other parameter could be used generically. We make no use of the properties of this variable, other than as an identifier of active region magnetism. In addition, we simplify the equations we use by considering only their overall magnitude, as opposed to their distribution in latitude. This allows one to understand the overall transfer

of flux in our view from the smallest scales visible (EPRs) to the largest scale fields (unipolar magnetic regions [UMRs]) of which the polar field is the clearest and most persistent example. In this section, we generate dynamo-type behavior using a simplified model, one without spatial structure, although spatial structure is required in a detailed model, as we did earlier. Although in the previous computations, we required stochastic processes, here with differential equations they present difficulties that we avoid owing to the form of our equations. To compensate for the lack of spatial structure in the differential equations, in the next section we discuss the model, in terms of the magnetic field's spatial structure and the numerical size of quantities. Let us add that the purpose of our simplification is not to fit the Sun well, but rather to illustrate the global effects of percolation. Thus we have also left out the important meridional flow and differential rotation, which are key elements to the Sun's real behavior. Fortunately, Schrijver (2001) has done an excellent job in his simulations of the Sun's dynamo, which include surface features and flow patterns. Our simulations remove many realistic aspects and thus their primary function is illustrative rather than total realism.

Let us begin with some basics, related to dynamo theory. In an MHD dynamo, the fluid behaves "as a fluid." Namely, the separate charged components of the gas do not move separately and thus cannot generate current, without pre-existing field being magnified. A battery behaves differently; it can generate current and magnetic field via the diverse motions of the electrons and ions, without any initial magnetic field. Hence for an MHD dynamo, magnetic field cannot be magnified without amplifying pre-existing field. Thus, for an MHD dynamo, one need consider how magnetic fields become twisted and magnified by the contorted motions of the conducting fluid, which can stretch loops of field into greater field strength by pulling on the field, thereby adding energy to this "rubber-band-like material," the magnetic field. For all but the first equation, each term has a "field growth term" leading to a transferring of one field (e.g., the sunspot number field, Rz), at the expense of another field term, a "shrinkage term" (e.g., the toroidal field, Bt). For the first equation, the EPR field grows from convective mixing of subscale fields, too small to see, gaining energy from the superadiabatic gradient, S . The subscale fields are operating on magnetic field, but this field is too small to measure. It first becomes measurable when the ephemeral regions form. In the remaining equations, the forms are basically nonlinear, with sets of pairs of variables leading to growth and shrinkage. They describe how the energy is drained from the superadiabatic gradient into the EPRs and subsequently into sunspots. The sunspot fields subsequently decompose into the larger scale structures of UMRs.

Figure 6 shows the overall flow of energy, via the conversion of magnetic field between varieties of features in the Sun. Field originates by convection associated with the superadiabatic gradient near the Sun's surface from fields too small to measure. They generate pores/ephemeral regions in the magnetic carpet. These grow from percolation into small, then larger sunspots gaining strength from the superadiabatic gradient. After this phase, magnetic fields are thought to distort into toroidal fields and poloidal fields, in accord with the Babcock-Leighton model. We will see (from the equations) that all three features are connected to and interact with each other in our model. This is shown by the double arrows in Figure 6 connecting each feature. Sunspots originate from percolation, but their configuration is determined by their interactions with the poloidal and toroidal fields (near the surface in this model). This interaction leads to Joy's law, with the geometry shown by Babcock. The magnetic flux in these processes results in an oscillatory process as the poloidal and sunspot fluxes

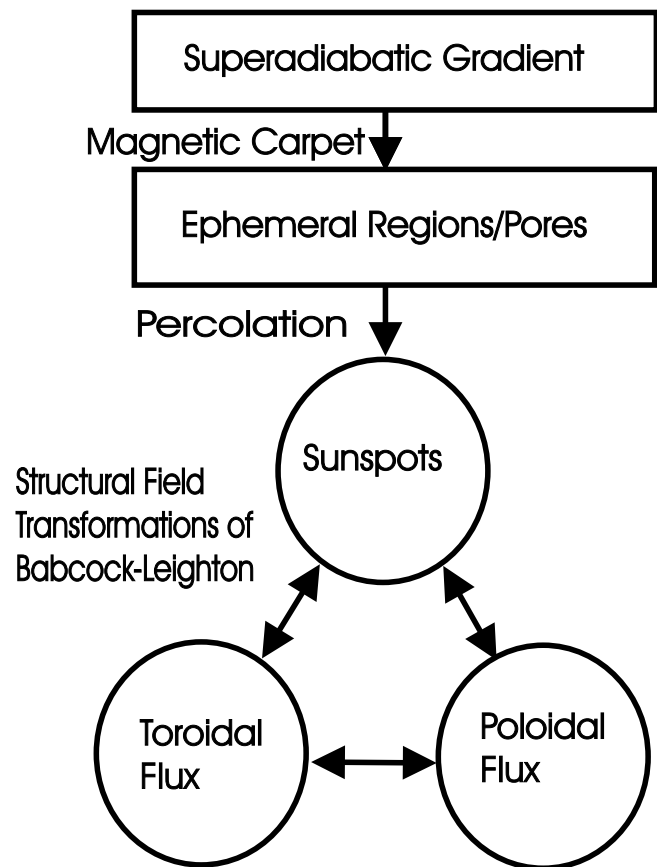


FIG. 6.—Illustrates the conversion of magnetic flux between different features in the photosphere, in accord to the current model. Magnetic flux is initiated by convection associated with the superadiabatic gradient near the Sun's surface from small fields. This generates pores/ephemeral regions in the magnetic carpet. These grow from percolation into sunspots. After this phase, magnetic fields are thought to distort into toroidal flux and poloidal flux, according to the Babcock-Leighton (B-L) model. We show double arrows, connecting all three features in the B-L model, as these all interact with each other. Sunspots originate from percolation, but their configuration is determined by their interactions with the poloidal and toroidal fields (near the surface in this model). This interaction leads to Joy's law, as Babcock showed.

compete for the available energy. This oscillation, generated by these equations, is thought to behave very much like the 11 year solar cycle. If our model is correct, then without the introduction of new magnetic flux from ephemeral regions, these field oscillation processes would decay, on the order of decades, suggesting the highly important role that ephemeral regions play for magnifying very weak fields into near kilogauss strength, as well as their known coronal role. It may be important to study the amount of flux in ephemeral regions, as a means of predicting solar activity on timescales longer than a decade.

Although Leighton (1964, 1969) provided a number of variations (e.g., deep vs. shallow dynamo models), his model had basically three equations (his eqs. [8]–[10]), in addition to initial condition equations. They governed temporal variations associated with the toroidal field, the radial field, and an unerupted field. They were one-dimensional with latitude as a spatial variable, calculated for μ (cosine latitude) variations of 0.1. Finer results were obtained by interpolation. To obtain a stochastic behavior, Leighton undertook random variations, which actually are somewhat problematic with differential equations, owing to the discrete and chaotic behavior of these two disparate forms. Nevertheless, he overcame this obstacle by employing terms that dampened growth, numerical noise, and other effects. We modify

similar equations to the ones he employed but modified in two ways: we remove the stochastic noise, and we integrate over latitude and thus only consider the average behavior, by representing each of the following terms by global values. Our model is based on the following four terms. We consider differential equations associated with the amount of flux, rather than just magnetic field, however, we think the meaning in the Leighton equations was similar. We consider the following magnetic entities: (1) an “insertion of amount of flux,” chosen to be the smallest features emanating into the photosphere, namely the number of ephemeral regions (EPRs), (2) an amount of flux gathered by the percolation process into larger scale (active region size) entities; here we arbitrarily choose sunspot number, Rz , to represent the larger scale field values although others might choose total sunspot area, or any other function of solar activity (the particular variable is really just a dummy variable), (3) the toroidal flux, ϕ_T and the poloidal flux, ϕ_P . The choice of sunspot number might seem deleterious to some, but those opposed to this parameter can substitute any other parameter as a representation for active region magnetic flux.

The EPRs are assumed to grow based on the local superadiabaticity, S . Although locally, S may vary, for our purposes here, we assume the global value does not vary much with activity, although observations will be needed to ascertain how much this varies. This energy source is assumed sufficiently powerful that the number of ephemeral regions is not depleted much by the negative feedback term associated with solar activity. In addition consider equations integrated over latitude, and to simplify all equations, we both ignore geometric factors (e.g., functions of cosine latitude, μ), and write each variable in dimensionless form, so that each variable is dimensionless (e.g., t is time/ T_0 , where T_0 is a period). Below we consider overall levels, so that our equations can be referenced to solar conditions:

$$\frac{dEPR}{dt} = f1(S). \quad (4.1)$$

Although the superadiabaticity, S , may grow near the Sun’s surface locally when sunspot activity ceases, and shrink toward zero and may go negative after an activity center depletes the available energy, when the growth of an active region ceases, for the global Sun as a latitudinal and time average, it is assumed approximately constant, $S = K$. When the secular trend of solar activity is constant, it implies that $f1(S) = 0$. A growth in solar activity implies $f1(S) > 0$, and vice versa.

The second equation governs how the number or amount of flux growing from EPRs into larger flux regions (e.g., sunspots) changes based on the number or magnetic flux in ephemeral regions, and how the regions are removed as sunspots interact with the toroidal field. Thus, the next equation has terms chosen that are functions of these three quantities. Although the two functional forms are undoubtedly complex, we approximate them by ignoring the detailed form and even a multiplying constant. This allows us to examine the solution without details of the numerical size or detailed shape. We write the equations with general functional forms, f_n , where n is the variously numbered terms, but then simplify the terms to particularly simple functions having some degree of nonlinear behavior, recognizing that actual processes on the Sun will be much more complex:

$$\frac{dRz}{dt} = f2(Rz, EPR) - f3(Rz, \phi_T) \sim RzEPR - Rz\phi_T. \quad (4.2)$$

In equations (4.2)–(4.4), we have chosen the simplest functional Lotka-Volterra forms that allow the following: self-similar

forms of growth (e.g., dynamo equations, so that magnetic fields are amplified by pre-existing magnetic field), and decay terms (using general functional forms, f_n). These forms, $f1$ – $f6$, allow any functional form of parameters, but then we start with linearity (of each form, but owing to the number of equations and multiplicity of parameters, i.e., xy , the overall behavior is nonlinear) as the simplest choice, with which to investigate the functional dependences. The interaction between the various forms provides a dynamo type behavior. If we had chosen higher order dependences, our curves would have differing shapes, but the essence, within a certain acceptable range, would be the same. Our choice of linearity within each term is an ad hoc choice, however, surprisingly for sunspots the arbitrary choice of linearity might not be too bad. Schrijver (2001) undertakes simulations of stellar dynamos growing from flux emergence and finds that the flux grows proportional to the injection rate to the power of 1.01 ± 0.01 . Thus, the choice of linearity seems to be supported by his study, although we have no further justification for this. It may be possible to modify the Schrijver model along the lines outlined here, to form a more realistic solar dynamo, essentially allowing the rate ephemeral region appearance to give rise to the rate of sunspot development.

The next equation governs toroidal flux growth. Babcock and Leighton had similar views in this regard, namely the growth of toroidal field occurring beneath the photosphere had its amplification initiated from a poloidal field that was increasingly wrapped by differential rotation. This feature is seen in Hale’s law of sunspot polarities, and as Babcock expresses it: “After 3 years, the equator will have gained in its rotation about 5.6 turns on the latitude circles at $\phi = \pm 55^\circ$,” Leighton had this term as $\partial''B_r/\partial t = -aB_\phi/(2\pi R\tau)$, where the constant evaluated to $\sim 18\sin^2\theta$ radians per year. For large spots this term became nonlinear as Leighton considered it as $\partial''B_r/\partial t = -a_0|B_\phi|B_\phi/(2\pi RB_c\tau)$, where B_c was a critical field that needed to be surpassed in order for the field to erupt. He also had exponential decaying terms, our first term, and the other terms are nonlinear in a manner similar to his terms (using two field parameters), with the difference being that we do not use the radial field term. Note that the positive signed nonlinear term in equation (4.3) is balanced by the negative term in equation (4.2), and so on for all the nonlinear terms. Again using the same “approximation,” the functions are simplified to yield the simpler form on the right:

$$\begin{aligned} \frac{d\phi_T}{dt} &= -\phi_T + f4(Rz, \phi_T) - f5(\phi_T, \phi_P) \\ &= -\phi_T + Rz\phi_T - \phi_T\phi_P. \end{aligned} \quad (4.3)$$

We have not explicitly written the influence of the Sun’s differential rotation in this approximation on the right of equation (4.3), as we are considering the main source of energy of the dynamo to be the superadiabatic gradient, which drives the fields in Rz . Although the form of differential rotation clearly is important in the spatial aspects of the Sun’s dynamo, we are only viewing the rough temporal variations here. Equation (4.3) might benefit by future inclusion of differential rotation explicitly, rather than just considering it a “geometrical factor” that we have ignored for simplicity. Equation (4.4) governs poloidal flux growth. Again we have an exponential decaying term, and a nonlinear growth term, similar to Leighton’s equations, again simplifying with our very rough approximation:

$$\frac{d\phi_P}{dt} = -\phi_P + f6(\phi_T, \phi_P) = -\phi_P + \phi_P\phi_T. \quad (4.4)$$

Figures 7 and 8 show two results from runs of these equations. In Figure 7, we simply have a steady state input. The figure shows

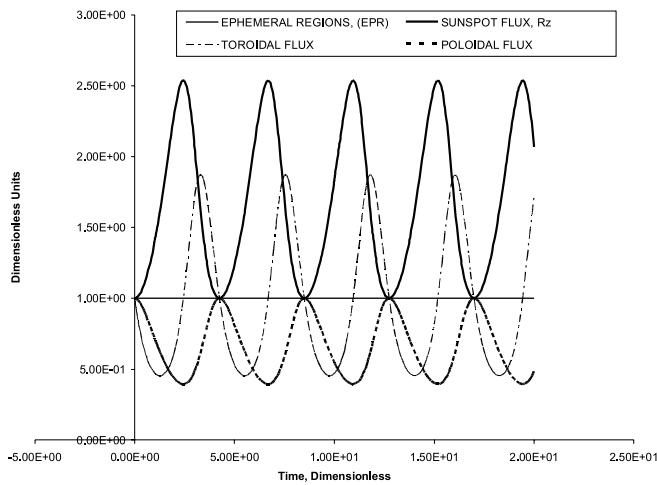


FIG. 7.—Shows the temporal variations of ephemeral regions (EPRs), sunspot flux, R_z , toroidal flux, and poloidal flux. In this simplest model, the ephemeral region input is constant. All parameters are in dimensionless units, meaning that they could be normalized to their mean values. The temporal oscillations develop from the nonlinear interaction of the four variables. Of interest is the out of phase behavior of the sunspot and poloidal fluxes (*heavy solid and heavy dashed lines, respectively*), as in the B-L models.

cyclic behavior develops by feeding each term off of the others, with the exception of the Ephemeral Regions, assumed constant, and a nonlinear oscillation develops. Of interest is how the poloidal flux and sunspot numbers, or flux, are out of phase, very much as in the solar cycle, with the polar fluxes maximizing near solar minimum. How would our differential equations behave if we introduced a small secular trend to our source term, the number of ephemeral regions?

Figure 8 displays this. The number of ephemeral regions goes up $\sim 30\%$ and then down. As can be seen from the nonlinearity, the sunspot number and poloidal flux values increase by a factor of 3 or 4. In addition, the periodicity is inversely correlated with the amplitude. These variations are similar to those seen between grand solar minima periods, like the Maunder minimum, and more active time periods. Waldmeier (1961, p. 171) found the same kind of inverse correlation between solar cycle amplitude and cycle length. The in-phase variation of the polar fluxes and the next cycle's sunspot numbers, seen in Figure 8, supports our use of polar fluxes for predicting the next cycle's sunspot number. This has been used (see Schatten et al. 1978; Svalgaard et al. 2005; Schatten 2005) to predict, fairly successfully, three past solar cycles. Prior to those predicted cycles, it was tested (see Schatten et al. 1972) with proxy data using eight solar cycles. The current model displays the kind of correlation that these authors have utilized. Now we discuss some of the numbers associated with these variables, and how they fit into the dynamo geometry.

5. DISPOSITION WITHIN THE SOLAR DYNAMO

In this section we discuss the placement of the current views within the framework of the solar dynamo, particularly the Babcock-Leighton viewpoint. It may help to understand this in terms of the field geometry shown in Figure 9. It modifies the Babcock picture panel A into two panels, below, which illustrate how magnetic fields change during the early and late phases of a solar cycle.

We view the essentials of the Babcock-Leighton (B-L) field geometry and their primary findings as unchanged. Most of the essential features of their model (e.g., the connection of field geometry with Hale's laws, the Spoerer butterfly diagram, Joy's

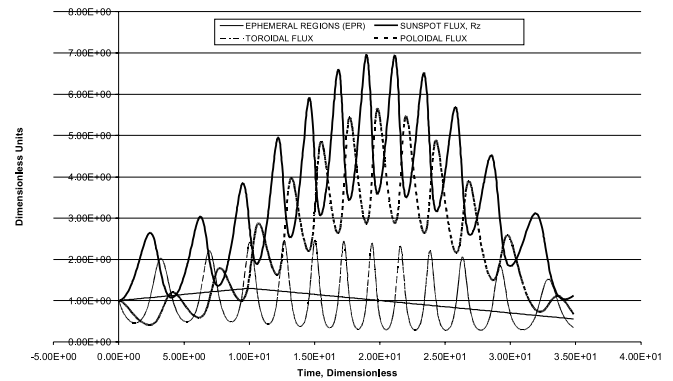


FIG. 8.—Similar to Fig. 7, however the input of ephemeral regions have a small initial growth, then decrease, rather than a constant value. Secular temporal variations of sunspot, toroidal and poloidal fluxes result. In addition, the cycles shorten as they grow, in accord with observations as found by Waldmeier (1961, p. 171). The figure also illustrates that poloidal flux and sunspot flux increase together (although 180° out of phase), tracking very well, thereby allowing the poloidal flux near solar minimum to serve as a precursor for future sunspot numbers.

law) remain well supported by observations. The main focus of this paper has been to consider the initiation of sunspot fields, really unknown at the time of these authors. Without the modern observations of ephemeral regions and recognition of the magnetic carpet, these authors could not have considered these fields as a source of the Sun's dynamo. We think that some representation of this type of shallow magnetic field source is in the spirit of the B-L shallow dynamo models, but not in the upwelling model that these models have evolved into. Observational support for their model can be found in a rather nice personal review by Sheeley (2005).

Let us step back to re-examine our overall picture of the Sun's dynamo. The Sun's convection zone first germinates high-frequency, small spatial scale magnetic fields (the observed ephemeral regions/pores), as part of the amplification of anomalies that occurs in a highly unstable, highly conducting environment transporting huge amounts of energy. For the Sun, this behavior serves essentially as a noise generator or amplifier magnifying small-scale fields from the granulation and supergranulation. These small field regions then combine or percolate into small spots, then group into larger spots and active regions through further percolation. As in Babcock's model, their location and direction is controlled by connections to the overall shallow field geometry (as supported by Hale's observations). As sunspots grow larger, they are further powered by convective energy transport, particularly downflows in the superadiabatic solar environment. They serve a role in helping make radially energy transport more efficient by essentially enforcing one-dimensional flow. This allows a greater efficiency (in transporting the Sun's energy outward) than the turbulent smaller scale convective overturning. A rather nonstandard model of Piddington's (1978) has some similarities to the present view. Although he seems to suggest a deep flux model, the accumulation of sunspot fields by accretion is also evident in his work. The depth of the origin of the photospheric magnetic flux and the multiplicity of field sources combining (the essence of percolation) are really two separate issues. We have not really distinguished these two separate issues within this paper, but have treated this view as indivisible throughout this paper. Nevertheless, we recognize there are two issues involving the birth of sunspots: that these magnetic fields (1) may arise from either shallow or deep layers, and (2) they may arise from the upwelling of a single flux tube or through the accumulation of many small

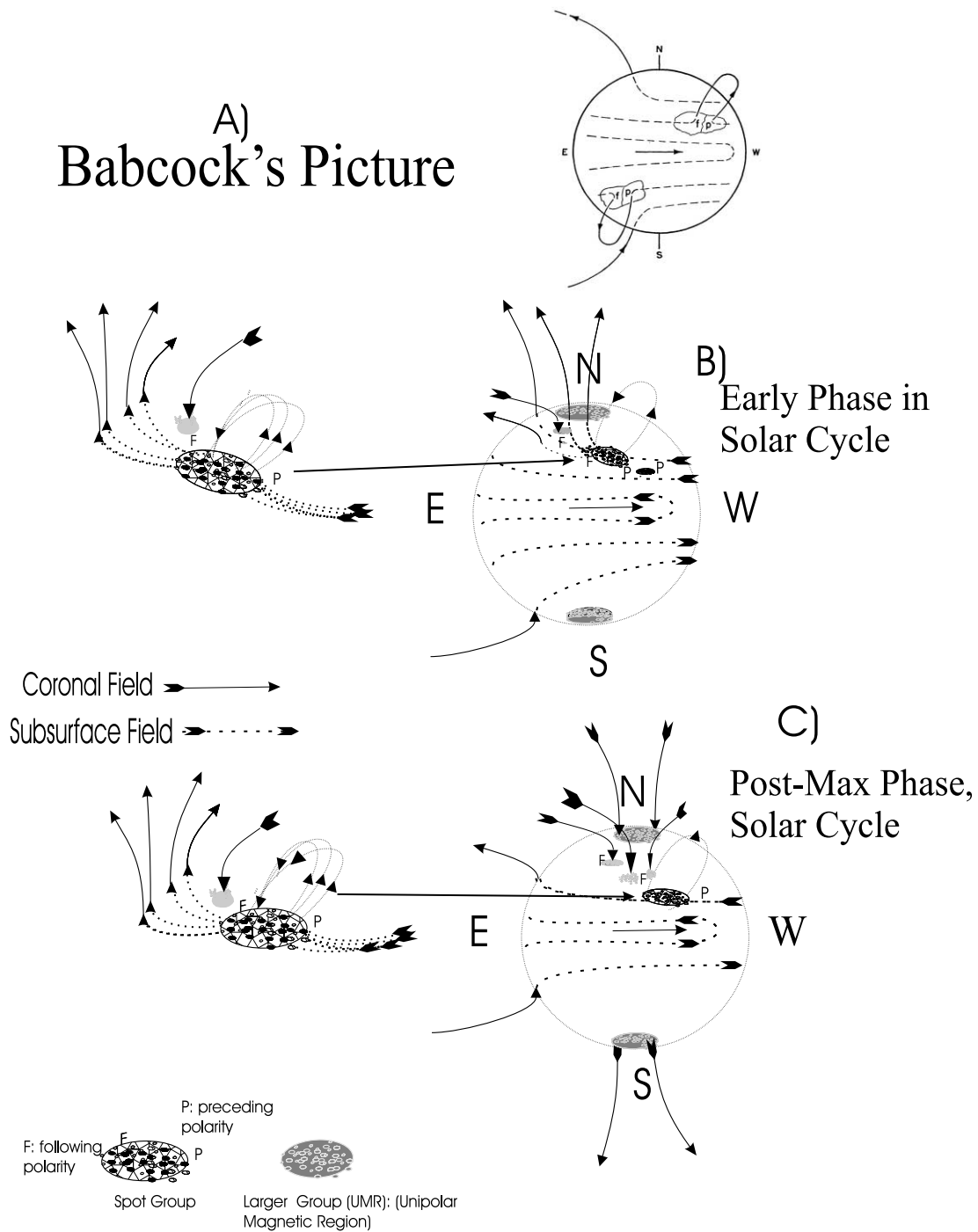


FIG. 9.— (a) Babcock's stage 3 geometry. The current view (b) shows magnetic field in the early phase of a solar cycle. The postmax phase is shown in (c). Active regions containing sunspots are shown by irregular shapes at low latitudes, with *p* and *f* referring to the preceding and following portions of the group. These generally contain opposite polarity. The relation of active region fields to the Sun's polar fields is given by Hale's laws. Panel *c* shows how the polar fields reverse near the solar maximum phase of a solar cycle. The low latitude spot groups are shown with darker shading, and the UMRs, including polar field, are shown by the light shading. *Solid lines*: Coronal field above the photosphere. *Dashed lines*: Magnetic field below the photosphere.

flux tubes via percolation. Nevertheless, in this paper, we have tended to associate a deep field origin with a single source for sunspots, and shallow field origins with multiple sources. It may be better to separate these two issues and examine them separately.

The percolation process may also be viewed as an “inverse” diffusion process, since it gathers together small fields, combining them into larger fields. Such an inverse process is behaving against the “normal manner” in which intensive thermodynamic

variables (temperature, magnetic field, pressure, etc.) tend to equalize. Thus, such inverse behavior is only possible in a nonequilibrium environment, such as the outer layers of the Sun's convection zone through the energy of convection that can do work. This allows like fields to be attracted to each other, rather than repel, as is their normal behavior in free space. The field separation can occur in accord with their interaction of the Sun's subsurface B-L field geometry (shown in Fig. 9), and this displays Joy's law,

as shown by Leighton. We see little reason for a deep source origin of Joy's law, when it seems to be available within the B-L framework.

Subsequent to the birth and development of active regions, the remainder of the solar dynamo is viewed in the standard Babcock-Leighton picture. The geometry of large-scale fields (unipolar regions, polar fields, large-scale sectors, etc.) may be viewed as a summation of Green's function solutions to the diffusion equation, possibly enhanced with meridional flow, driven by the sunspots as delta function sources. The Leighton model for the diffusive spread of spot fields was observationally tested directly (see Schatten et al. 1972). Although there was support for the Leighton model, there were also some puzzling aspects that Sheeley (2005) and Wang & Sheeley (1994) considered. Their model explained the rigid rotation of large-scale features stemming from differentially rotating sources. Their explanation is a nice addition to understanding how rigidly rotating patterns of field can develop from a differentially rotating field source. We view their considered patterns as "eigenmodes" that the field forms from the various differentially rotating inputs. Schulz (2001) has studied these aspects related to the coronal magnetic field, with diffusion along similar lines. We view both as essentially correct, the former just more mathematical, and the latter more physical. These all seem like different expressions for the same UMRs that Bumba found exist on the Sun. We view the UMRs as magnetic eigenmodes of the Sun's magnetic field that are "excited" or filled by the remnant fluxes from sunspot fields after the superadiabatic gradient has energized or enhanced the sunspot field values, and then these fields subsequently decay (the higher order eigenmodes quickly dampening out) into the low-order modes. Higher order effects may be considered by this diffusion model, as suggested by Sheeley (2005), by considering meridional and other flows, which would lead to an advection diffusion model, which has an adaptive grid model in Fortran for finding solutions to such models as advective diffusion physics.¹

Particular solutions of the large-scale field of the Sun may be obtained, after the methods of Schatten et al. (1972), using essentially a Green's function technique. We outline this for the diffusion equation, without advection, differential rotation, etc. but discuss the changes these aspects would require. The technique follows Morse & Feshbach (1953) and Jackson (1975). Taking a diffusion equation, that for heat transport, is written as

$$\frac{\partial \psi}{\partial t} = \alpha \nabla^2 \psi + \frac{\dot{q}}{\rho C_p}, \quad (5.1)$$

where ψ is temperature, α is the thermal diffusivity, \dot{q} is a release of thermal energy, ρ is density, and C_p , specific heat content in the heat equation. In our case ψ is the signed (+ or -) magnetic flux. In general this equation can be attacked via Green's function techniques. These work by finding general solutions to a linear equation (the first two terms in eq. [5.1]), with a hypothetical localized function in space near $r = r_0$ and time near $t = t_0$ replacing the heat input, thus allowing a summation of linear methods to work. One then applies the general solution as an integral over the discrete solutions. Taking this approach, the diffusion equation is then written as

$$\frac{\partial G}{\partial t} - \alpha \nabla^2 G = 4\pi \delta(r - r_0) \delta(t - t_0), \quad (5.2)$$

where G is the required Green's function. A solution to this equation may be obtained by separation of variables, so G can be written as

$$G(r - r_0, t - t_0) = 4\pi \alpha \left[\frac{1}{2\sqrt{\pi \alpha (t - t_0)}} \right]^n \exp \left[-\frac{(r - r_0)^2}{\alpha} \right] \quad (5.3)$$

in n dimensions. Placed into a dimensionless framework, the equation is often written in terms of functions of $\xi = r/2(\alpha t)^{1/2}$, with the solution being $T(r, t) = f(\xi)$, and removing all the constants and parameters in equation (5.2). The solution is then

$$f(\xi) = A \int \exp(-\xi^2) d\xi + B = A \frac{\sqrt{\pi}}{2} \operatorname{erf}(\xi) + B. \quad (5.4)$$

The familiar aspect of this solution is the term of distance divided by the square root of time, often arises in these type equations. It is familiar in diffusion, probability theory, kinetic theory, and many other areas of physics.

In our case, we are interested in diffusion of magnetic fields on a spherical shell, so let us focus in on that. For situations of a spherical shell, the geometry is often expanded into spherical harmonics, so that the spatial expansion is written in terms of angular functions, with a delta function put into form of

$$\delta(r - r_0) = \frac{1}{r^2} \delta(r - r_0) \delta(\phi - \phi_0) \delta(\cos \theta - \cos \theta_0), \quad (5.5)$$

with the angular dependence written in terms of spherical harmonics, this is

$$\delta(r - r_0) = \frac{1}{r^2} \delta(r - r_0) \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}^*(\theta_0, \phi_0) Y_{lm}(\theta, \phi). \quad (5.6)$$

The various terms then decay away with times depending on the diffusion coefficient, which Leighton estimates (of order $10^4 \text{ km}^2 \text{ s}^{-1}$). This coefficient provides a good value for the lifetime of UMRs, based on Bumba's observations.

The solution for a number of sources, M_i at various times, t_i , is then just

$$B(r, t) = \int_{-\infty}^t \sum_i M_i G(r - r_i, t - t_i) dt'. \quad (5.7)$$

For Green's functions including advective terms, e.g., differential rotation, one may consider the diffusion solutions (the spherical harmonic terms), to be differentially distorted from a framework of spherical trapezoidal structures (which are the nulls of the solutions) to ones where the sides are differentially rotated, as in Bumba's UMR shapes. Thus, various contributions would blend into a single value (but not necessarily the same at either pole). This provides a valuable method for estimating polar field evolution from active region sources, and a means to understand why high-latitude active regions (early in the cycle) play a more dominant role in reversing and forming a new cycle's polar field value than the lower latitude field sources late in a cycle.

Let us consider the values of some of the numbers associated with the equations and field parameters in the above equations. These values are patterned after the Babcock-Leighton models and the ephemeral regions, based on the observations of Parnell (2002). She estimates the average emergence rate for the magnetic carpet to be between 6×10^{-2} and $10^{-5} \text{ Mx cm}^{-2} \text{ s}^{-1}$. In

¹ See Adaptive Grid Refinement in Fortran, AGRIF, online at <http://www-lmc.imag.fr/IDOPT/AGRIF/index.html>.

accordance with her estimates, a magnetic field in sunspots to ~ 500 G would seem to require between ~ 0.1 and 50 days from a weak field via the separation and accumulation of magnetic carpet flux. Babcock takes the polar field to have $\sim 8 \times 10^{21}$ Mx, consisting of eight ropes, each rope defined as 10^{21} Mx. This provides ~ 3000 spot groups or BMRs, with each unit length providing 3–4 BMRs. With the Leighton (1964) model, a diffusion coefficient, D , of order $10^4 \text{ km}^2 \text{ s}^{-1}$ was used. This provides a time constant of order $\tau \sim \pi R_{\text{Sun}}^2 / D \sim 5$ yr, comparable to $\sim 1/2$ the solar cycle period of 11 years. Thus, in this model, the timescale is of order of the effective diffusion timescale for the large-scale field, the UMRs, which must make their way to the poles, in order to recycle flux.

From this, and from Sheeley and Wang's studies of the large-scale field, we suggest that the whole dynamo mechanism may be viewed as a high-frequency and small spatial scale "noise generator," of magnetic field (ephemeral regions and sunspots), which then is filtered by diffusion processes on the Sun (acting as a low-frequency, large spatial scale filter), allowing the longer timescale eigenmodes of unipolar magnetic regions to dominate and then gradually decay. The polar fields, being the longest lived eigenmodes, are then the largest scale dominant low field structures we see. The 11 year cycle has an associated timescale of $\sim (2\pi/22 \text{ yr})$, whose period is governed by the timescale for field growth, diffusion, reversal, etc., to completely reverse the magnetic field twice. This timescale, in our view, is just the time it takes for all these processes to work together, and thus is tuned at all, as seen in Figure 8, so no well defined period is strictly set. Hence changes in the magnitude of the field, or variations in the patterns of field on the Sun's surface, may affect the timing, as Waldmeier (1961, p. 171) and others report. Let us now turn to how this model may be tested.

6. OBSERVATIONAL TESTS

We provide observational tests that can be used to differentiate the current model from the conventional deep-seated, alpha-omega dynamo.

6.1. Test 1—Temporal Development of Activity Centers

To summarize, we expect to see in our model the development of activity centers gradually form from the accumulation of numerous small regions. In deep-seated dynamos, our understanding is that fluxes of equal and opposite sign would appear in the photosphere, virtually fully formed, as they rise rapidly to the surface, under the influence of ~ 27 times the Earth's gravity present in the solar atmosphere.

In the alpha-omega dynamo, to initiate a sunspot group, field from a flux tube erupts due to instabilities and buoyancy effects from the solar interior and rises rapidly to the Sun's photosphere. It is our understanding that magnetic fields from such an eruption should then have approximately equal positive (outward) and negative (inward) fluxes. The present model suggests a gradual growth from many small regions to larger ones through flux accumulation, in a manner similar to how storms accumulate moisture to form larger systems. Thus, we expect a gradual evolution from small ephemeral regions to larger regions that eventually form a sunspot group. Growth would be enhanced in regions where the superadiabatic gradient, S , is large. In addition, the field line picture of Babcock's, shown in Figure 9, suggests a fertile ground for active region growth exists in the fertile regions located equatorward and eastward, along the field direction suggested by Joy's law, from previous eruptions of magnetic flux.

The temporal development of the three-dimensional nature of solar magnetism and magnetic flux accumulation might be obtainable from high-resolution high-cadence observations of growing active regions from recent spacecraft such as *SOHO* MDI or modern ground-based solar telescopes. Of course such telescopes may not probe deeply into the Sun. Nevertheless, they may show same-sign flux from ephemeral regions growing into larger scale features in accord with the current view. Growth of active regions happens more frequently during the rising portion of an active solar cycle, so it is hoped that solar cycle 24 may serve as an opportune time to test this and other solar cycle theories.

6.2. Test 2—Magnetic Flux Distribution within Groups

The magnetic flux rising from a deep-seated flux tube would be expected to have a distribution of equal magnitudes of equal and opposite flux signs. This may be represented by a distribution function:

$$P(\phi) = N_p \delta(\phi - \phi_0) + N_N \delta(\phi + \phi_0), \quad (6.1)$$

so that a single flux tube erupting releases equal amounts of positive, $N_p \phi_0$, and negative, $-N_N \phi_0$, fluxes.

For our percolation model, flux should accumulate from surrounding areas. To represent this, a formula may be written as

$$P(\phi) = \sum_{i=\Delta \text{Area}} P(\phi - \phi_i) \Delta A, \quad (6.2)$$

where fluxes of various signs and magnitudes (ϕ_i may be + or –, and of any magnitude associated with some overall distribution possibly varying as the superadiabatic gradient) may accumulate in the active region area, as fluxes percolate, or accumulate in this region. This is suggested by equation (6.2). We recognize that field strengths may change magnitude (e.g., increase), as fields are compressed together, but for the sake of simplicity, we have not included that in equation (6.2).

The manner of flux accumulation in this model may appear complex, as many regions are involved. Nevertheless, it is simply many similar processes being repeated. Ephemeral regions should start with initial equal and opposite fluxes and may form in the presence of a surrounding environment of one sign field or the other. Then as many of these regions develop, they draw in surrounding surface fluid, containing their fluxes. If an active region is developing at some site, it is doing so because it is able to eject surface fluid downward, due to its cold, dense nature relative to other areas of the photosphere that are marginally hotter and lighter. The accumulating center would contain more field of one sign or the other, depending on the sign of the magnetic field within that environment.

For example, Bumba & Howard (1965) have identified the weak fields of the Sun as UMRs. If an active region develops in a UMR of one sign, it should have a larger amount of that sign flux, with the discrepancy growing with time (as opposite field is ejected), until the region decays. To clarify this, let us state it more explicitly. If an active region develops in an area with a positive polarity, for example, and that is the preceding polarity, then that sunspot area should predominate over the following spot areas (and vice versa). This suggests fairly straightforward tests of a difference between surface models (whether or not our mechanism is right or not) and deep source models for bipolar field initiation. Distributions based on equation (6.1) versus equation (6.2) compared with the surrounding flux distribution should hopefully provide some tests distinguishing the two models.

In addition, there have also been reported a number of unipolar spots (only one of a pair of the conventional dipole). The current hypothesis suggests that these would appear in strong UMRs, with the field having the polarity of the surrounding UMR. Presumably the field of the UMR and the bipolar fields would balance each other, in terms of net flux, so some kind of large-scale accounting might also be possible.

6.3. Test 3—Flow Patterns

The following flow pattern test might show some effects, although it seems less definitive than the previous two. For young active regions, inflows would be associated with flux magnification. Conversely, for old regions, outflows would be associated with flux decay. In addition, for young regions, same-sign flux of small field regions (e.g., ephemeral active regions, pores, small spots) would be attracted to that sign flux within the active region. So, the detailed motions within, or near, active regions could be studied. The motion is opposite to the conventional wisdom that magnetic flux rises from the solar interior. Rather than diverging from a growing region, magnetic field (of the proper polarity—following flux in a following region, etc.) and fluid should be converging toward the parts of a growing region. Decaying regions would show reverse motion, and the sunspots may change from overall energy deficits, as spots eject bright faculae (seen near the limb), as suggested by the observations of Harvey & Harvey (1973) and Hagenaar & Shine (2005).

7. SUMMARY AND CONCLUSIONS

We suggest a new method for sunspot initiation within the framework of the Babcock-Leighton solar dynamo models. The model starts with ephemeral regions (EPRs), which form the magnetic carpet, studied by Schrijver & Title (2001) and others at Lockheed. We then choose a seemingly unusual behavior for these small bipoles, based on Parker's theoretical work that fibril fields lower the total energy within a highly conducting convective flow environment (e.g., the upper convection zone).

This strange behavior may be briefly understood as follows. Conventionally, as in bar magnets, we are familiar with same sign, or "like," magnetic fields repelling. In addition, we understand that magnetic fields fill all of space. Yet on the surface of the Sun, surprisingly, the field sometimes behaves in a somewhat different manner. Overall, magnetic fields on the Sun are distributed into large areas of near-zero field, but small areas of very intense field on the order of a kilogauss (as in ephemeral regions, pores and sunspots). This dichotomy has been explained by Parker (1984), who showed that in a convective environment, the overall convective plus field energy is lowered by dividing space up into two disparate (high field, of one sign or the other, and near-zero field) regions. Essentially, this allows the majority of the solar atmosphere to transport energy through convection via uninhibited field-free regions, as the magnetic field would otherwise dampen the transport of energy, and hence raise the total energy.

In accord with this behavior, this paper posits the existence of a percolation force that essentially adds "stickiness" to same-sign magnetic field in areas ripe for field growth (namely, where the superadiabatic gradient is large), thereby allowing like fields to attract and thus behave as Parker suggests. Rather than calculating the dynamical motion of the magnetic fields subject to all the forces on it as might be possible with a few objects in Keplerian motion, we provide a cellular automaton approach, by utilizing two-dimensional cellular automata, such that like-sign fields stick. We refer to the growth of such features as percolation, a term that has been used in other fields similarly. Our

percolation model illustrates that small field structures can grow into larger ones, such as sunspots; however, our model does not fit the geometry of the bipolar nature of active regions. Percolation behavior for the Sun was also investigated by Seiden and Wentzel, who fit magnetic field distributions and did obtain a bipolar field configuration. This behavior, we suggest, allows a shallow solar dynamo with surface percolation to replace the birth of active regions through flux eruptions from the solar interior. We note that Babcock originally suggested a shallow solar dynamo and Leighton analyzed both a shallow and deep dynamo formation; most other features of the B-L dynamo remain. Thus, we are primarily reconsidering the birthplace of active regions from the interior to the photosphere. We further suggest, however, that rather than the differential rotation supplying much of the dynamo energy, it is the superadiabatic gradient (the outflowing solar energy) that drives the Sun's dynamo. This may be supported by the Sun issuing forth more energy, when solar activity is large (more energy is transported outward). Differential rotation and meridional flow are still important for the detailed structure and development of spot groups.

We also support our model through the development and numerical simulation of differential equations that illustrate the temporal behavior of a dynamo model. These are a form of simplified Leighton equations that allow their basic nature to be examined. We also provide observational tests that can be used to differentiate our model from the conventional alpha-omega dynamo. In particular, we suggest three observational tests for this model:

1. The temporal development of activity centers should grow from many small features gradually accumulating in size and number, rather than being born anew as a complete entity.

2. The magnetic flux distribution within activity centers in our model bears a relationship to the surrounding distribution of flux, prior to that group's development. In the current model, fields of one sign would be more likely to grow from regions containing that sign flux. Thus, amid a unipolar region of one sign (e.g., positive), we might expect more of that sign polarity (positive) in the sunspots within that region, as opposite flux is ejected.

3. Flow pattern tests are expected as follows, in this model. Young active regions are expected to have inflows and be associated with flux magnification. Conversely, for old regions, outflows would be associated with ejection of same-sign magnetic flux and hence field decay. Detailed motions within, or near, active regions could be studied, and are expected to yield same sign fluxes attracting as regions grow. Rather than diverging from a growing region, magnetic field (of the proper sign) and fluid should be converging. Decaying regions would show reverse motion, and the sunspots may change from energy deficit, dark regions to bright faculae (seen near the limb).

A word of caution is also advised about finding some evidence of solar activity in deep solar layers implying a deeply seated dynamo. Surface phenomena (e.g., cool downflows associated with activity belts) may affect deeper layers and serve as a tracer of recent activity. This would not necessarily indicate that the source of activity is in the deep layers. The effect is similar to a plane dropping leaflets to the ground. It can leave evidence on the ground of the leaflets, as well as their proper motion, yet the plane itself is not on the ground.

From Sheeley and Wang's studies of the large-scale field, we suggest that the whole dynamo mechanism may be viewed as a high-frequency, small spatial scale "noise generator," of magnetic field (ephemeral regions and sunspots), which is then filtered by diffusion processes on the Sun (acting as a low-frequency, large spatial scale filter). This allows the longer timescale eigenmodes

of unipolar magnetic regions (UMRs) and polar fields to dominate the large-scale magnetic structures. While the present model is highly simplified, which we acknowledge, we also suggest that there is some strength in having the model's internal workings "transparent." With the current model being predominantly illustrative, it is envisaged that more realistic shallow solar dynamo models will be forthcoming, thereby allowing others to improve the model, as they see fit. Finally, our modeling suggests that the amount of flux generated in ephemeral regions, if it can be accurately monitored for long periods, may possibly lead to long

timescale (multidecadal) predictions of solar activity. Solar activity variations may show a nonlinear dependence on the amount of flux generated in ephemeral regions.

The author appreciates comments from Hans Mayr, Eugene Parker, Leif Svalgaard, and Donat Wentzel. In addition, the author appreciates the many helpful critical comments by an anonymous referee.

REFERENCES

- Adami, C. 1999, *Introduction to Artificial Life* (New York: Springer)
- Babcock, H. W. 1961, *ApJ*, 133, 572
- Brandenburg, A. 2005, *ApJ*, 625, 539
- Bumba, V., & Howard, R. F. 1965, *ApJ*, 141, 1502
- Chapman, G. A., Herzog, A. D., & Lawrence, J. K. 1986, *Nature*, 319, 654
- Cox, J. P., & Giuli, R. T. 1968, *Principles of Stellar Structure* (New York: Gordon & Breach)
- Dikpati, M., de Toma, G., & Gilman, P. A. 2006, *Geophys. Res. Lett.*, 33, 5102
- Hagenaar, H. J., & Shine, R. J. 2005, *ApJ*, 635, 659
- Harvey, K. L. 1993, Ph.D. thesis, Univ. Utrecht
- Harvey, K. L., & Harvey, J. 1973, *Sol. Phys.*, 28, 61
- Hurlburt, N. E., & Toomre, J. 1988, *ApJ*, 327, 920
- Jackson, J. D. 1975, *Classical Electrodynamics* (2nd ed.; New York: John Wiley & Sons)
- Lean, J. L., Livingston, W. C., Heath, D. F., Donnelly, R. F., Skumanich, A., & White, O. R. 1982, *J. Geophys. Res.*, 87, 10307
- Leighton, R. B. 1964, *ApJ*, 140, 1547
- . 1969, *ApJ*, 156, 1
- Lin, H., & Rimmele, T. 1999, *ApJ*, 514, 448
- Longcope, D., & Choudhuri, A. 2002, *Sol. Phys.*, 205, 63
- Meyer, P. 2000, Masters thesis, Univ. Derby
- Mihalas, D. 1970, *Stellar Atmospheres* (San Francisco: Freeman)
- Morse, P. M., & H. Feshbach. 1953, *Methods of Theoretical Physics* (New York: McGraw-Hill)
- Nandy, D., & Choudhuri, A. R. 2002, *Science*, 296, 1671
- . 1979, *ApJ*, 232, 291
- . 1984, *ApJ*, 283, 343
- . 2001, *Chinese J. Astron. Astrophys.*, 1, 99
- Parnell, C. 2002, *Astron. & Geophys.*, 43, 4.16
- Piddington, J. H. 1978, *Ap&SS*, 55, 401
- Press, W. H., Flannery, B. P., & Teukolsky, S. A. 1986, *Numerical Recipes in Fortran: The Art of Scientific Computing* (Cambridge: Cambridge Univ. Press)
- Ruzmaikin, A. 2001, *Space Sci. Rev.*, 95, 43
- Schatten, K. H. 1988, *ApJ*, 329, 1028
- . 2005, *Geophys. Res. Lett.*, 32, 21106
- Schatten, K. H., Leighton, R. B., Howard, R., & Wilcox, J. M. 1972, *Sol. Phys.*, 26, 283
- Schatten, K. H., & Mayr, H. G. 1985, *ApJ*, 299, 1051
- Schatten, K. H., Scherrer, P. H., Svalgaard, L., & Wilcox, J. M. 1978, *Geophys. Res. Lett.*, 5, 411
- Schrijver, C. J. 2001, *ApJ*, 547, 475
- Schrijver, C. J., & Title, A. M. 2001, *ApJ*, 551, 1099
- Schulz, M. 2001, *J. Geophys. Res.*, 106, A8, 15859
- Seiden, P. E., & Wentzel, D. G. 1996, *ApJ*, 460, 522
- Sheeley, N. R., Jr. 2005, *Living Rev. Sol. Phys.*, 2, <http://www.livingreviews.org/lrsp-2005-5>
- Sofia, S., Oster, L., & Schatten, K. 1982, *Sol. Phys.*, 80, 87
- Stauffer, D. 1994, *Introduction to Percolation Theory* (New York: Plenum Press)
- Svalgaard, L., Cliver, E. W., & Kamide, Y. 2005, *Geophys. Res. Lett.*, 32, L01104
- Svalgaard, L., & Wilcox, J. M. 1976, *Sol. Phys.*, 49, 177
- Waldmeier, M. 1961, *The Sunspot-Activity in the Years 1610–1960* (Zurich: Schulthess)
- Wang, Y.-M., & Sheeley, N. R. Jr., 1994, *ApJ*, 430, 399
- Webster, E. W. 2004, Aristotle's "Meteorology" (Adelaide: Univ. Adelaide), <http://etext.library.adelaide.edu.au/a/aristotle/meteorology/complete.html>
- Willson, R. C., Gulkis, S., Janssen, M., Hudson, H. S., & Chapman, G. A. 1981, *Science*, 211, 700